



TUGAS AKHIR - SS141501

**PEMODELAN WAKTU *SURVIVAL* PEKERJA  
MENGUNAKAN REGRESI *COX PROPORTIONAL*  
*HAZARD***

**NADIA RAHMA NUR INSYIRA**  
**NRP 062114 4000 0080**

**Dosen Pembimbing**  
**Dr. R. Mohamad Atok, S. Si., M.Si**  
**Imam Safawi Ahmad S.Si., M. Si.**

**PROGRAM STUDI SARJANA**  
**DEPARTEMEN STATISTIKA**  
**FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA**  
**INSTITUT TEKNOLOGI SEPULUH NOPEMBER**  
**SURABAYA 2018**





**TUGAS AKHIR - SS141501**

**PEMODELAN WAKTU SURVIVAL PEKERJA  
MENGUNAKAN REGRESI *COX PROPORTIONAL*  
HAZARD**

**NADIA RAHMA NUR INSYIRA  
NRP 062114 4000 0080**

**Dosen Pembimbing  
Dr. R. Mohamad Atok, S. Si., M.Si  
Imam Safawi Ahmad S.Si., M. Si.**

**PROGRAM STUDI SARJANA  
DEPARTEMEN STATISTIKA  
FAKULTAS MATEMATIKA, KOMPUTASI, DAN SAINS DATA  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2018**

*“Halaman ini sengaja dikosongkan”*



**FINAL PROJECT - SS141501**

**MODELING OF SURVIVAL TIME PERFORMANCE  
USING REGRESSION COX PROPORTIONAL HAZARD**

**NADIA RAHMA NUR INSYIRA  
SN 062114 4000 0080**

**Supervisors**

**Dr. R. Mohamad Atok, S. Si., M.Si  
Imam Safawi Ahmad S.Si., M. Si.**

**UNDERGRADUATE PROGRAMME  
DEPARTMENT OF STATISTICS  
FACULTY OF MATHEMATICS, COMPUTING, AND DATA SCIENCE  
INSTITUT TEKNOLOGI SEPULUH NOPEMBER  
SURABAYA 2018**

*“Halaman ini sengaja dikosongkan”*

## LEMBAR PENGESAHAN

### PEMODELAN WAKTU *SURVIVAL* PEKERJA MENGUNAKAN REGRESI *COX PROPORTIONAL* *HAZARD*

#### TUGAS AKHIR

Diajukan Untuk Memenuhi Salah Satu Syarat  
Memperoleh Gelar Sarjana Sains  
pada  
Program Studi Sarjana Departemen Statistika  
Fakultas Matematika, Komputasi, dan Sains Data  
Institut Teknologi Sepuluh Nopember

Oleh:

**Nadia Rahma Nur Insyira**

NRP. 062114 4000 0080

Disetujui Oleh Pembimbing:

**Dr. R. Mohamad Atok, S. Si., M.Si**

NIP. 19710915 199702 1 001

**Imam Safawi Ahmad S.Si., M. Si.**

NIP. 19810224 201404 1 001

()  
Imam w f o r m a  
-



Mengetahui,  
Kepala Departemen

  
**Dr. Suhartono**

DEPARTEMEN  
STATISTIKA

NIP. 19710929 199512 1 001

SURABAYA, JULI 2018

*“Halaman ini sengaja dikosongkan”*



# **PEMODELAN WAKTU *SURVIVAL* PEKERJA MENGUNAKAN REGRESI *COX PROPORTIONAL* HAZARD**

**Nama Mahasiswa** : Nadia Rahma Nur Insyira  
**NRP** : 062114 4000 0080  
**Departemen** : Statistika  
**Dosen Pembimbing** : Dr. R. Mohamad Atok, S. Si., M.Si  
Imam Safawi Ahmad S.Si., M. Si.

## **ABSTRAK**

*Kematian merupakan peristiwa yang tidak dapat diperkirakan waktu kejadiannya sehingga dapat dialami oleh semua orang. Kematian dapat disebabkan oleh berbagai macam faktor, dalam penelitian ini diduga ada lima faktor yang mempengaruhi kematian pekerja atau aparatur sipil negara (ASN). Penelitian ini bertujuan untuk mengetahui model survival yang terjadi pada kematian pekerja aparatur sipil negara (ASN). Faktor-faktor yang diduga mempengaruhi kematian seorang pekerja akan dimodelkan dengan waktu ketahanan hidup (waktu survival), sehingga akan diketahui faktor-faktor mana saja yang berpengaruh secara signifikan terhadap waktu survival pekerja. Penelitian ini akan dilakukan pendekatan dengan metode bootstrap, yang merupakan pengambilan sampel secara random berdasarkan data asli sehingga akan didapatkan parameter-parameter yang signifikan dengan melihat nilai confidence interval. Berdasarkan hasil analisis Kaplan-Meier untuk semua variabel menurun sampai rentang waktu 40 tahun. Pengujian log-rank diketahui bahwa pada variabel jenis kelamin dan anak tidak ada perbedaan secara signifikan antara kurva survival disetiap kelompoknya. Berbeda dengan variabel golongan pekerja dan pasangan diketahui bahwa variabel tersebut memiliki perbedaan kurva survival antar kelompoknya. Pemodelan regresi cox proportional hazard menghasilkan variabel golongan dan pasangan yang berpengaruh secara signifikan terhadap waktu survival pekerja.*

**Kata Kunci :** *Kematian, Kaplan-Meier, Log-Rank, Pekerja, Regresi Cox Proportional Hazard*

*“Halaman ini sengaja dikosongkan”*

## **MODELING OF SURVIVAL TIME PERFORMANCE USING REGRESSION COX PROPORTIONAL HAZARD**

**Name** : Nadia Rahma Nur Insyira  
**Student Number** : 062114 4000 0080  
**Department** : Statistics  
**Supervisors** : Dr. R. Mohamad Atok, S. Si., M.Si  
Imam Safawi Ahmad S.Si., M. Si.

### **ABSTRACT**

*Death is an event that cannot be predicted, so everyone's death can be happened so suddenly. Death can be caused by various factors, in this research allegedly there are five factors which caused the death of a worker or a civil state apparatus. The goal of this research is to know the survival model which happened in the death of a worker or a civil state apparatus. The alleged factors that cause a worker's death will be modeled with survival time, so it can be known which factors that affect the worker's survival time significantly. In this research the bootstrap method will be used in the approach, to take samples randomly based on the original data, so that it can generate the parameters significantly. Based on the Kaplan-Meier analysis of all variables decreased to a range of 40 years old. Meanwhile, from the log-rank test it can be concluded that the survival curves of variables gender and children are not significantly different in each group. Different from the variables of the worker category and partner, it can be concluded that those two variables have different survival curves in each group. The cox proportional hazard regression modelling yield that variable partner has a significant effect towards a worker's survival time.*

**Keyword** : *Death, Kaplan-Meier, Log-Rank, Regression Cox Proportional Hazard, Worker*

*“Halaman ini sengaja dikosongkan”*

## KATA PENGANTAR

Puji syukur kehadirat Allah SWT atas limpahan rahmat dan hidayahNya sehingga penulis dapat menyelesaikan Tugas Akhir yang berjudul **“Pemodelan Waktu *Survival* Pekerja dengan Menggunakan Regresi *Cox Proportional Hazard*”** dengan baik. Penulis menyadari bahwa penyusunan Tugas Akhir ini tidak terlepas dari dukungan dan bantuan dari berbagai pihak. Oleh karena itu, penulis mengucapkan terima kasih kepada:

1. Bapak Dr. Suhartono selaku Kepala Departemen Statistika yang telah memberikan banyak fasilitas untuk kelancaran penyelesaian Tugas Akhir ini.
2. Bapak Dr. R. Mohamad Atok, M.Si dan Bapak Imam Safawi Ahmad S.Si., M. Si. selaku dosen pembimbing yang telah memberikan nasihat, kritik, saran dan waktu kepada penulis dalam menyelesaikan Tugas Akhir ini.
3. Bapak Dr. Ir. Setiawan, M.S. dan Ibu Santi Puteri Rahayu, M.Si., Ph.D. selaku dosen penguji atas saran dan kritiknya yang sangat membangun.
4. Bapak Dr. Sutikno, M.Si. selaku dosen wali atas nasihat dan saran yang telah diberikan.
5. Orang tua, adik, dan seluruh keluarga dari ibu penulis yang selalu mendukung dengan segala doa, kasih sayang, nasihat, dan dukungan yang diberikan kepada penulis.
6. Teman-teman terdekat Mas Eky, anak kos cendrawasih, teman-teman seperjuangan Respect 2014  $\Sigma 25$ , teman-teman SMA, teman-teman WO dan teman-teman penulis lainnya.
7. Semua pihak yang mendukung dalam penyelesaian Tugas Akhir Besar harapan penulis untuk mendapatkan kritik dan saran yang membangun sehingga Tugas Akhir ini dapat bermanfaat bagi semua pihak yang terkait.

Surabaya,     Juli 2018  
   Penulis

*“Halaman ini sengaja dikosongkan”*

## DAFTAR ISI

	Halaman
<b>HALAMAN JUDUL</b> .....	i
<b>LEMBAR PENGESAHAN</b> .....	v
<b>ABSTRAK</b> .....	vii
<b>ABSTRACT</b> .....	ix
<b>KATA PENGANTAR</b> .....	xi
<b>DAFTAR ISI</b> .....	xiii
<b>DAFTAR GAMBAR</b> .....	xvii
<b>DAFTAR TABEL</b> .....	xix
<b>DAFTAR LAMPIRAN</b> .....	xxi
<b>BAB I PENDAHULUAN</b> .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	4
1.3 Tujuan Penelitian .....	4
1.4 Manfaat Penelitian .....	5
1.5 Batasan Masalah.....	5
<b>BAB II TINJAUAN PUSTAKA</b> .....	7
2.1 Analisis <i>Survival</i> .....	7
2.1.1 Tipe Penyensoran Data.....	8
2.1.2 Fungsi <i>Survival</i> dan Fungsi <i>Hazard</i> .....	10
2.2 Kurva <i>Survival</i> Kaplan-Meier dan Uji <i>Log-Rank</i> .....	13
2.3 Regresi Cox .....	15
2.3.1 Asumsi <i>Proportional Hazard</i> .....	15
2.3.2 Model Regresi <i>Cox Proportional Hazard</i> .....	16
2.3.3 Estimasi Parameter Model <i>Cox Proportional Hazard</i> .....	17
2.3.4 Seleksi Model Terbaik.....	19
2.3.5 Pengujian Signifikansi Parameter .....	20
2.3.6 <i>Hazard Ratio</i> .....	21
2.3.7 <i>Bootstrap</i> .....	21
2.4 Kematian .....	23
<b>BAB III METODOLOGI PENELITIAN</b> .....	25
3.1 Sumber Data.....	25
3.2 Variabel Penelitian .....	25

3.3	Langkah-langkah Analisis Data .....	27
<b>BAB IV</b>	<b>ANALISIS DAN PEMBAHASAN .....</b>	<b>31</b>
4.1	Analisis Karakteristik Variabel pada Data Pekerja .....	31
4.1.1	Karakteristik Waktu <i>Survival</i> ( $T$ ) dan Usia ( $X_5$ )..	31
4.1.2	Karakteristik Jenis Kelamin ( $X_1$ ) .....	32
4.1.3	Karakteristik Golongan Pekerja ( $X_2$ ) .....	34
4.1.4	Karakteristik Pasangan ( $X_3$ ).....	36
4.1.5	Karakteristik Anak ( $X_4$ ).....	38
4.2	Kurva Kaplan-Meier dan Pengujian <i>Log-Rank</i> untuk Semua Variabel Independen yang Mempengaruhi <i>Survival Time</i> .....	40
4.2.1	Karakteristik Pekerja Berdasarkan Semua Variabel Independen.....	40
4.2.2	Karakteristik Pekerja Berdasarkan Variabel Kategorik Jenis Kelamin ( $X_1$ ) dan Uji <i>Log-Rank</i> .....	41
4.2.3	Karakteristik Pekerja Berdasarkan Variabel Kategorik Golongan Pekerja ( $X_2$ ) dan Uji <i>Log- Rank</i> .....	43
4.2.4	Karakteristik Pekerja Berdasarkan Variabel Kategorik Pasangan ( $X_3$ ) dan Uji <i>Log-Rank</i> .....	44
4.2.5	Karakteristik Pekerja Berdasarkan Variabel Kategorik Anak ( $X_4$ ) dan Uji <i>Log-Rank</i> .....	45
4.3	Pemodelan Regresi <i>Cox Proportional Hazard</i> .....	46
4.3.1	Pengujian Asumsi <i>Proportional Hazard</i> .....	46
4.3.2	Estimasi Parameter Regresi <i>Cox Proportional Hazard</i> .....	48
4.3.3	Seleksi Model Terbaik Regresi <i>Cox Proportional Hazard</i> .....	49
4.3.4	Interpretasi Model <i>Cox Proportional Hazard</i> .....	51
4.3.5	Analisis Regresi <i>Cox Proportional Hazard</i> dengan Menggunakan Metode <i>Bootstrap</i> .....	52
<b>BAB V</b>	<b>KESIMPULAN DAN SARAN .....</b>	<b>55</b>
5.1	Kesimpulan .....	55
5.2	Saran.....	56



<b>DAFTAR PUSTAKA .....</b>	<b>57</b>
<b>LAMPIRAN .....</b>	<b>59</b>
<b>BIODATA PENULIS .....</b>	<b>679</b>

*“Halaman ini sengaja dikosongkan”*

## DAFTAR GAMBAR

<b>Gambar 2.1</b> Data Sensor Tipe I .....	8
<b>Gambar 2.2</b> Data Sensor Tipe II .....	9
<b>Gambar 2.3</b> Data Sensor Tipe III .....	10
<b>Gambar 2.4</b> Ilustrasi Kurva Kaplan Meier.....	16
<b>Gambar 3.1</b> Diagram Alir .....	28
<b>Gambar 3.2</b> Diagram Alir Lanjutan.....	29
<b>Gambar 4.1</b> <i>Pie Chart</i> Jenis Kelamin .....	32
<b>Gambar 4.2</b> <i>Pie Chart</i> Golongan Pekerja .....	34
<b>Gambar 4.3</b> <i>Pie Chart</i> Pasangan.....	36
<b>Gambar 4.4</b> <i>Pie Chart</i> Anak .....	38
<b>Gambar 4.5</b> Kurva Kaplan-Meier Semua Variabel Independen .....	41
<b>Gambar 4.6</b> Kurva Kaplan-Meier untuk Variabel Jenis Kelamin .....	42
<b>Gambar 4.7</b> Kurva Kaplan-Meier untuk Golongan Pekerja .....	43
<b>Gambar 4.8</b> Kurva Kaplan-Meier untuk Variabel Pasangan .....	44
<b>Gambar 4.9</b> Kurva Kaplan-Meier untuk Variabel Anak.....	45

*“Halaman ini sengaja dikosongkan”*

## DAFTAR TABEL

<b>Tabel 3.1</b>	Variabel Dependen Penelitian .....	25
<b>Tabel 3.2</b>	Variabel Independen Penelitian.....	26
<b>Tabel 3.3</b>	Struktur Data Penelitian.....	26
<b>Tabel 4.1</b>	Statistika Deskriptif untuk Data Pekerja.....	31
<b>Tabel 4.2.</b>	Tabulasi Silang Jenis Kelamin dengan Status Pekerja.....	32
<b>Tabel 4.3</b>	Tabulasi Silang Rata-Rata Waktu <i>Survival</i> dan Usia pada Variabel Jenis Kelamin.....	33
<b>Tabel 4.4.</b>	Tabulasi Silang Status Pekerja dengan Golongan .....	35
<b>Tabel 4.5.</b>	Tabulasi Silang Rata-Rata Waktu <i>Survival</i> dan Usia pada Variabel Golongan Pekerja.....	35
<b>Tabel 4.6.</b>	Tabulasi Silang Status Pekerja dengan Pasangan .....	36
<b>Tabel 4.7.</b>	Tabulasi Silang Rata-Rata Waktu <i>Survival</i> dan Usia pada Variabel Pasangan .....	37
<b>Tabel 4.8.</b>	Tabulasi Silang Status Pekerja dengan Anak .....	39
<b>Tabel 4.9.</b>	Tabulasi Silang Rata-Rata Waktu <i>Survival</i> dan Usia pada Variabel Anak.....	39
<b>Tabel 4.10</b>	Hasil <i>Goodness of Fit</i> Semua Variabel.....	47
<b>Tabel 4.11</b>	Hasil <i>Goodness of Fit</i> Tanpa Variabel Jenis Kelamin .....	47
<b>Tabel 4.12</b>	Estimasi Parameter Model <i>Cox Proportional</i> <i>Hazard</i> .....	48
<b>Tabel 4.13</b>	Estimasi Parameter Model <i>Cox Proportional</i> <i>Hazard</i> Terbaik .....	50
<b>Tabel 4.14</b>	<i>Confidence Interval</i> 95% dari Parameter dengan Metode <i>Bootstrap</i> .....	52

*“Halaman ini sengaja dikosongkan”*

## DAFTAR LAMPIRAN

<b>Lampiran 1</b>	Data Pekerja Pensiun Meninggal Aktif .....	59
<b>Lampiran 2</b>	<i>Output</i> SPSS Uji <i>Log-Rank</i> Jenis Kelamin .....	60
<b>Lampiran 3</b>	<i>Output</i> R <i>Goodness of Fit</i> .....	61
<b>Lampiran 4</b>	Estimasi Parameter Seluruh Variabel yang Signifikan.....	62
<b>Lampiran 5.</b>	Seleksi Model Terbaik dengan Kriteria AIC .....	63
<b>Lampiran 6.</b>	Estimasi Parameter Model Terbaik untuk Regresi <i>Cox Proportional Hazard</i> .....	64
<b>Lampiran 7</b>	Syntax Regresi <i>Cox Proportional Hazard</i> dengan Metode <i>Bootstrap</i> .....	65
<b>Lampiran 8</b>	Surat Legalitas Data .....	67

*“Halaman ini sengaja dikosongkan”*



# **BAB I**

## **PENDAHULUAN**

### **1.1 Latar Belakang**

Kematian merupakan peristiwa yang tidak dapat diperkirakan waktu kejadiannya, hal ini dapat dialami oleh semua orang. Kematian itu sendiri memiliki arti berakhirnya fungsi biologis tertentu. Contohnya hilangnya pernafasan, berhentinya aliran tekanan darah dan tubuh menjadi kaku. Hal-hal ini menjadi tanda-tanda seseorang dianggap mengalami peristiwa kematian (Santrock, 2002). Tidak hanya orang yang berusia lanjut tetapi dapat dialami oleh bayi, balita, remaja, dan orang dewasa.

Penelitian ini akan diteliti tentang faktor-faktor penyebab kematian pada data pekerja Pegawai Negeri Sipil. Data ini tercatat sebagai data pensiun meninggal aktif di PT. Taspen Kantor Cabang Utama Surabaya. Data pensiun meninggal aktif merupakan data yang dicatat ketika seorang keluarga peserta pensiun melakukan klaim meninggal sebelum memasuki waktu pensiun. Faktor-faktor yang diduga mempengaruhi kematian pekerja adalah jenis kelamin pekerja, golongan pekerja, pasangan pekerja, anak pekerja, dan usia pekerja sehingga variabel yang digunakan dalam penelitian ini akan dibagi menjadi dua bagian yaitu variabel dependen dan variabel independen. Variabel dependen dalam penelitian ini adalah lama seseorang bekerja sampai meninggal atau masa kerja seorang pekerja. Variabel independennya adalah jenis kelamin pekerja, golongan pekerja, pasangan pekerja, anak pekerja dan usia. Variabel-variabel tersebut digunakan untuk mencari faktor mana saja yang mempengaruhi pekerja meninggal. Penelitian ini data akan dibagi menjadi dua tipe yaitu data tersensor dan data tidak tersensor, dimana data tersensor pada penelitian ini merupakan pekerja yang mengalami kematian sebelum bulan Desember tahun 2017.

Terpilihnya variabel jenis kelamin sebagai salah satu pendugaan yang mempengaruhi pekerja meninggal karena berdasarkan data yang sudah diterbitkan oleh *World Health Organization* (WHO) usia angka harapan hidup wanita lebih

panjang daripada pria. Usia angka harapan hidup wanita pada data yang diterbitkan WHO sebesar 71,2 tahun dan usia angka harapan hidup pria sebesar 67,1 tahun. Selain itu terdapat studi yang dilakukan oleh *Institute of Health Metrics and Evaluation (IHME)* di *University of Wahington* yang melibatkan 188 negara selama 1990-2013. Menurut studi ini usia harapan hidup penduduk Indonesia untuk wanita berada di urutan ke-120 dan pria di urutan ke-106. Sedangkan dalam survei yang dilakukan Badan Pusat Statistika (BPS) angka harapan hidup wanita menunjukkan lebih tinggi daripada pria. Rata-rata nasional angka harapan hidup wanita berada pada angka 73 tahun sedangkan laki-laki berada pada angka 69 tahun.

Faktor lainnya yang diduga menyebabkan kematian adalah tuntutan pekerjaan yang diwakili dalam variabel golongan pekerjaan. Tuntutan pekerjaan yang tinggi dapat menyebabkan stres yang berkepanjangan dan menimbulkan gangguan kesehatan. Hal ini ditunjukkan oleh tim peneliti dari *Indiana University Kelley School of Business* yang menemukan bahwa jika seseorang berada di tempat kerja yang memiliki tuntutan pekerjaan tinggi maka resiko meninggal akan lebih besar 15,4% daripada yang tidak. Contoh gangguan kesehatan yang dapat menimbulkan kematian karena stres adalah stroke, jantung, dan darah tinggi (Gonzalez-Mule & Cockburn, 2016).

Jumlah anggota keluarga yang terlalu banyak diduga dapat menjadi faktor penyebab kematian. WHO memiliki riset yang menjadi bukti bahwa keluarga yang memiliki anak banyak, resiko kematian ibu semakin besar. Selain itu terdapat riset yang dilakukan oleh Maria Chiu dari *Institute for Clinical Evaluative Sciences* di Toronto yang dimuat di jurnal *the Lancet* tentang resiko kematian yang dialami oleh ayah tunggal. Jurnal tersebut mengatakan bahwa resiko ayah tunggal meninggal tiga kali lipat lebih tinggi daripada ibu tunggal dan yang berpasangan (Chiu, Rahman, & Vigod, 2018).

Berdasarkan sumber-sumber tersebut akan dibuktikan apakah faktor-faktor tersebut mempengaruhi kematian atau tidak

dengan menggunakan pemodelan waktu *survival* dari lama seseorang bekerja sampai meninggal atau masa kerja pekerja. Penelitian ini dilakukan dengan tujuan memberikan gambaran bagi perusahaan untuk mengetahui faktor-faktor mana saja yang mempengaruhi waktu *survival* dari pekerja. Metode yang cocok untuk digunakan dalam kasus ini adalah pemodelan regresi *cox proportional hazard* dikarenakan data yang diberikan oleh PT. Taspen merupakan data pengamatan waktu dengan banyak variabel-variabel independennya yang mempengaruhi variabel dependennya. Metode ini dapat memodelkan data *survival* dengan variabel penjelas yang mempengaruhi fungsi *hazard*. Model regresi *cox* dibuat oleh Sir David Cox. Kelebihan dari metode ini adalah dalam membentuk model tidak diperlukan fungsi distribusi parametrik dan asumsi pemodelan hanya untuk memvalidasi asumsi bahwa fungsi *hazard* harus proporsional setiap waktu.

Selanjutnya akan dibentuk model baru dengan mensimulasikan data asli dengan menggunakan metode *bootstrap*. Metode ini digunakan untuk mengatasi jumlah sampel yang sedikit dan dikhawatirkan hasil dari parameter-parameter pemodelan tersebut bias. Pada simulasi metode ini akan dilakukan pengambilan data secara acak dari data asli dengan mengulangnya sebanyak 5000 kali. Hal ini dilakukan untuk mendapatkan *confidence interval* dari setiap parameternya. Kelebihan dari metode ini adalah dapat digunakan untuk ukuran data yang relatif kecil dan lebih baik dari metode tradisional lainnya seperti bayes karena metode *bootstrap* menghasilkan panjang interval yang lebih pendek.

Penelitian sebelumnya dengan menggunakan regresi *cox proportional hazard* telah diteliti oleh beberapa orang. Salah satunya adalah Apriyanti (2018) yang meneliti pemodelan ketahanan hidup penderita hemodialisis di RSUD Kabupaten Sidoarjo dengan menggunakan metode regresi *cox proportional hazard*. Model tersebut digunakan untuk mengetahui faktor-faktor yang mempengaruhi transplantasi ginjal. Selain itu, penelitian yang dilakukan oleh Fa'rifah (2012) menggunakan metode regresi *cox*

untuk meneliti faktor-faktor yang mempengaruhi laju kesembuhan pasien penderita Demam Berdarah *Dengue* (DBD) di rumah sakit umum haji. Sedangkan untuk data pensiun meninggal aktif pernah diteliti oleh Anindita (2017) untuk dihitung hak dan manfaat yang diterima oleh ahli waris bagi peserta yang meninggal aktif. Data ini diperoleh dari PT. Taspen (Persero) Kantor Cabang Bandar Lampung.

Unzet dan Sanches (1992) telah meneliti tentang penerapan *bootstrap* pada analisis *survival* dengan menggunakan Kaplan-Meier. Penelitian tersebut menghasilkan kesimpulan bahwa metode *bootstrap* pada analisis *survival* dapat mengestimasi mediannya. Selain itu, penelitian yang dilakukan oleh Azimmatul Ihwah (2012) dengan menggunakan metode *bootstrap* membuktikan bahwa metode ini dapat mengestimasi koefisien beta pada model regresi *cox proportional hazard*. Hal ini menunjukkan bahwa *resampling bootstrap* dapat diterapkan pada data *survival*. Penelitian lain yang mengkaji *resampling bootstrap* adalah penelitian Rahayu (2012), dari lima variabel hanya terdapat tiga variabel yang signifikan berpengaruh terhadap data ketahanan hidup pasien.

## **1.2 Rumusan Masalah**

Penelitian ini didasari oleh banyaknya pekerja yang meninggal sebelum masa pensiun. Penelitian ini menggunakan data yang diberikan oleh PT. Taspen Kantor Cabang Utama Surabaya tentang data pensiun meninggal aktif. Data pensiun meninggal aktif ini akan diolah untuk menentukan faktor-faktor apa saja yang mempengaruhi kematian pekerja sehingga dapat diketahui model terbaik yang signifikan untuk permasalahan ini.

## **1.3 Tujuan Penelitian**

Berdasarkan rumusan permasalahan tersebut, tujuan yang ingin dicapai untuk penelitian ini adalah sebagai berikut :

1. Dapat mendeskripsikan karakteristik pekerja berdasarkan faktor-faktor yang diduga mempengaruhi waktu survival pekerja.

2. Memperoleh hasil identifikasi kurva survival pekerja menurut variabel independennya secara umum berdasarkan analisis Kaplan-Meier dan uji Log-Rank.
3. Mengetahui faktor-faktor yang signifikan dengan model regresi *cox proportional hazard* untuk data *survival* pekerja dari PT. Taspen Kantor Cabang Utama Surabaya dan mensimulasikannya dengan *resampling bootstrap*.

#### **1.4 Manfaat Penelitian**

Penelitian ini diharapkan dapat memberikan beberapa manfaat bagi berbagai pihak, diantaranya sebagai berikut:

1. Memberikan informasi bagi PT. Taspen (Persero) mengenai faktor-faktor mana saja yang paling berpengaruh dalam kematian seorang peserta pensiun.
2. Memberikan wawasan tentang statistika mengenai penerapan analisis survival dan regresi *cox proportional hazard* pada bidang asuransi.

#### **1.5 Batasan Masalah**

Batasan masalah digunakan untuk memfokuskan permasalahan yang akan diselesaikan. Batasan-batasan masalah yang digunakan dalam penelitian ini adalah sebagai berikut :

1. Data yang digunakan merupakan data pensiun meninggal aktif pada Pegawai Negeri Sipil (PNS) yang berada di PT. Taspen Kantor Cabang Utama Surabaya.
2. Periode data yang diambil bulan Februari 2018 sampai Maret 2018.

*“Halaman ini sengaja dikosongkan”*

## BAB II TINJAUAN PUSTAKA

Pada bagian tinjauan pustaka ini akan dibahas tentang metode-metode yang akan digunakan dalam menganalisis 100 data pekerja. Metode yang akan dibahas dalam bab ini adalah analisis *survival*, kurva Kaplan-Meier dan uji *Log-Rank*, analisis regresi *cox proportional hazard*, dan analisis regresi *cox proportional hazard* dengan metode *bootstrap*.

### 2.1 Analisis Survival

Analisis *survival* mempunyai tujuan menduga probabilitas kelangsungan hidup, kekambuhan, kematian, dan peristiwa-peristiwa lainnya sampai pada periode waktu tertentu. Tiga elemen yang harus diperhatikan dalam waktu *survival* adalah sebagai berikut (Le, 1997):

- a. *Time origin or starting point* (titik awal) adalah waktu dimulainya pekerja tercatat sebagai Pegawai Negeri Sipil di PT. Taspen.
- b. *Ending event of interest* (kejadian akhir) adalah kejadian dimana pekerja meninggal
- c. *Measurement scale for the passage of time* (skala ukuran untuk berlalunya waktu).

Perbedaan dari analisis *survival* dengan analisis statistik lainnya adalah analisis *survival* memiliki data tersensor. Data tersensor merupakan pengamatan waktu *survival* hanya sebagian, tidak sampai waktu kejadian. Fungsi sebaran kumulatif untuk peluang kejadian gagal sebelum waktu jangka di awal adalah sebagai berikut (Lee, 1980) :

$$F(y) = P(Y \leq y) = \int_0^y f(y) dt \quad (2.1)$$

Dengan

$Y$  : waktu *survival*

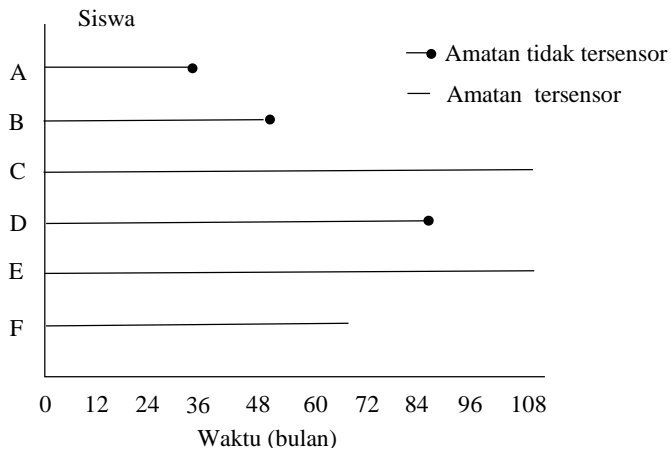
$f(y)$  : fungsi kepadatan peluang

### 2.1.1 Tipe Penyensoran Data

Data tersensor adalah keadaan dimana obyek penelitian keluar dari penelitian dikarenakan waktu hidup obyek penelitian tidak diketahui sampai terjadi sesuatu kejadian yang tidak terduga (Collet, 2003). Data tersensor dibagi menjadi 3 jenis, yakni :

#### 1. Data Sensor Tipe I

Data tipe ini merupakan percobaan dimana dilakukan selama waktu  $T$  yang telah ditentukan sampai berakhir waktu  $T$ . Hal ini dapat dinyatakan sebagai waktu tersensor. Pada tipe ini semua obyek yang diuji  $n$  masuk pada waktu yang sama dan tidak ada obyek yang hilang dalam waktu penelitian sehingga waktu tahan hidup observasi tersensor adalah waktu selama observasi berlangsung. Berikut ini contoh data sensor tipe I dengan menggunakan waktu ketahanan siswa.



**Gambar 2.1** Data Sensor Tipe I

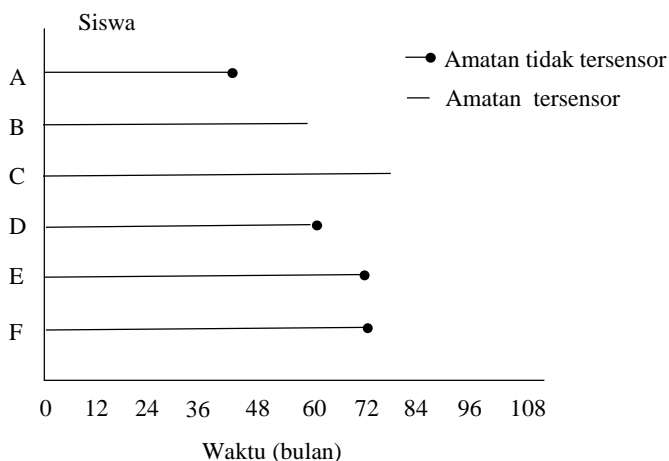
Berdasarkan Gambar 2.1 diketahui bahwa siswa A,B, dan D putus sekolah pada bulan ke-36, 48 dan 84. Siswa A, B, dan D ini merupakan obyek tidak tersensor. Siswa lainnya merupakan obyek tersensor. Siswa C dan E merupakan contoh kasus dimana obyek tetap bertahan hingga batas



penyensoran (*withdrawn alive*) akan tetapi siswa F merupakan kasus obyek hilang sebelum batas waktu penyensoran (*lost to follow up*).

2. Data Sensor Tipe II

Data tipe ini merupakan data yang diperoleh setelah waktu observasi mencapai kegagalan sebanyak  $r$  dari  $n$  data observasi yang ada. Pada tipe ini semua obyek yang diuji  $n$  masuk pada waktu yang sama dan jika ada obyek yang hilang maka waktu tahan hidup observasi sama dengan waktu tahan hidup observasi tidak tersensor terbesar. Berikut ini, contoh dari data sensor tipe II.



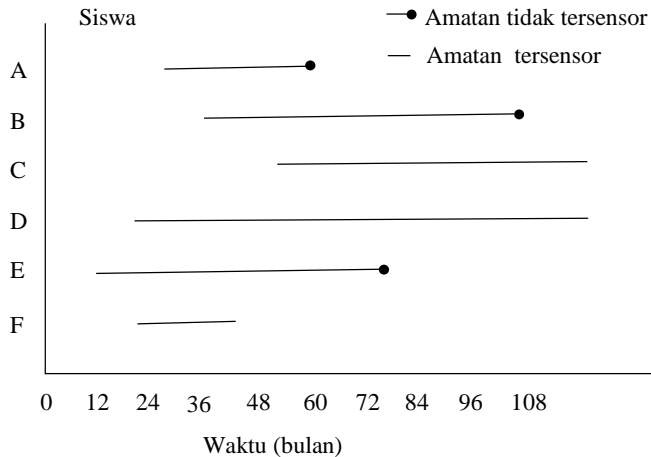
**Gambar 2.2** Data Sensor Tipe II

Gambar 2.2 menunjukkan data waktu ketahanan sekolah dari siswa SD. Penelitian akan dihentikan jika terdapat empat siswa yang putus sekolah. Berdasarkan gambar tersebut penelitian diakhir pada bulan ke-72 karena terdapat empat siswa yang putus sekolah yaitu siswa A, D, E dan F.

3. Data Sensor Tipe III

Data ini merupakan tipe data yang obyek penelitiannya masuk percobaan dalam waktu yang berbeda tetapi berhenti pada waktu yang ditentukan oleh peneliti. Contoh dari data

sensor tipe III merupakan data waktu ketahanan sekolah pada siswa.



**Gambar 2.3** Data Sensor Tipe III

Penelitian ini dilakukan dalam periode 9 tahun dimana terdapat enam siswa yang masuk kedalam pengamatan. Informasi ini dirujuk dalam Gambar 2.3. Pada gambar tersebut siswa A, B dan E masuk ke dalam pengamatan pada bulan ke-24, 36 dan 12. Maka dapat disimpulkan siswa A, B dan E merupakan siswa putus sekolah. Namun siswa C dan D merupakan kasus *withdrawn alive* dan siswa F merupakan kasus *lost to follow up*. Kasus *lost to follow up* tidak ada dalam penelitian ini karena peneliti tidak mengamati secara langsung obyek penelitian dari waktu awal pengamatan ketahanan bersekolah.

### 2.1.2 Fungsi *Survival* dan Fungsi *Hazard*

Pada penelitian ini terdapat dua fungsi untuk mengetahui probabilitas waktu *survival* pekerja dari *start point* sampai waktu  $t$  (Lee & Wang, 2003). Fungsi distribusi peluang  $f(t)$  merupakan fungsi dari variabel *random* dan waktu *survival* pada notasi  $T$ .  $T$  merupakan variabel *random* yang non negati dalam interval  $[0, \infty)$

sehingga *probability density function* dapat dinyatakan sebagai berikut :

$$f(t) = \lim_{\Delta t \rightarrow 0} \frac{P(t \leq T < t + \Delta t)}{\Delta t} \quad (2.2)$$

*Cumulative distribution function* merupakan fungsi yang dapat didefinisikan dengan peluang suatu obyek mengalami kejadian sampai dengan waktu  $t$  yang dapat dinyatakan sebagai berikut :

$$F(t) = P(T \leq t) = \int_0^t f(t) dt \quad (2.3)$$

Probabilitas suatu obyek atau dalam kasus ini adalah pekerja bertahan setelah waktu ke- $t$  dapat dinyatakan dengan persamaan fungsi *survival*  $S(t)$  sebagai berikut :

$$S(t) = P(T > t) = 1 - F(t) = 1 - P(T \leq t) \quad (2.4)$$

Fungsi lain yang akan dipakai pada penelitian ini adalah fungsi *hazard*. Fungsi *hazard* merupakan laju kegagalan (*failure*) dari suatu obyek penelitian yang mengalami kejadian (*event*) pada waktu ke- $t$  atau dapat dikatakan sebagai peluang obyek mengalami suatu kejadian dalam waktu ke- $t$ . Persamaan dari fungsi ini adalah sebagai berikut :

$$h(t) = \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t \mid T \geq t)}{\Delta t} \right\} \quad (2.5)$$

Fungsi *hazard* merupakan laju kegagalan obyek penelitian ketika mengalami event dalam rentang waktu  $t$  sampai dengan  $t + \Delta t$  dengan syarat obyek penelitian tersebut telah bertahan sampai waktu ke- $t$ . Hubungan antara fungsi *survival* dan fungsi *hazard* adalah teori probabilitas bersyarat untuk  $P(A \cap B)$ .  $P(A \cap B)$  merupakan probabilitas kejadian bersama antara  $A$  dan  $B$  sehingga probabilitas bersyarat dapat dirumuskan sebagai berikut :

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Dari persamaan 2.2 didapatkan bahwa jika dimisalkan  $f(t)$  adalah *probability density function* pada waktu  $t$ , maka didapatkan persamaan sebagai berikut :

$$\begin{aligned}
 h(t) &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < t + \Delta t | T \geq t)}{\Delta t} \right\} \quad (2.6) \\
 &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < (t + \Delta t) \cap (T \geq t))}{\Delta t \times P(T \geq t)} \right\} \\
 &= \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < (t + \Delta t))}{\Delta t \times S(t)} \right\} \\
 &= \frac{1}{S(t)} \times \lim_{\Delta t \rightarrow 0} \left\{ \frac{P(t \leq T < (t + \Delta t))}{\Delta t} \right\} \\
 &= \frac{1}{S(t)} \times \lim_{\Delta t \rightarrow 0} \frac{F(t + \Delta t) - F(t)}{\Delta t}
 \end{aligned}$$

Hubungan antara fungsi *survival* dan fungsi *hazard* dapat dinyatakan sebagai berikut :

$$h(t) = \frac{f(t)}{S(t)} \quad (2.7)$$

Berdasarkan persamaan (2.7), maka didapatkan persamaan sebagai berikut :

$$f(t) = \frac{d(F(t))}{dt} = \frac{d(1 - S(t))}{dt} = -\frac{d(S(t))}{dt} \quad (2.8)$$

$$h(t) = \frac{f(t)}{S(t)} = -\frac{dS(t)}{dt} \cdot \frac{d \ln S(t)}{dS(t)} = -\frac{d \ln S(t)}{dt} \quad (2.9)$$

Dari persamaan (2.9), dapat diperoleh persamaan sebagai berikut :

$$\begin{aligned}
 \int_0^t h(t) dt &= -\int_0^t \frac{d \ln S(t)}{dt} dt \\
 -\int_0^t h(t) dt &= \int_0^t \frac{d \ln S(t)}{dt} dt
 \end{aligned}$$

maka,

$$-\int_0^t h(t) dt = \ln S(t) \Big|_0^t = \ln S(t) - \ln S(0) \quad (2.10)$$

Karena  $S(0) = 1$ , maka  $\ln S(0) = \ln 1 = 0$ . Oleh karena itu, persamaan 2.10 dapat dituliskan menjadi sebagai berikut :

$$\begin{aligned} -\int_0^t h(t) dt &= \ln S(t) \\ \exp\left(-\int_0^t h(t) dt\right) &= \exp(\ln S(t)) \\ S(t) &= \exp\left(-\int_0^t h(t) dt\right) \end{aligned}$$

Sehingga, fungsi *survival* dapat dirumuskan pada persamaan 2.10.

$$S(t) = \exp(-H(t))$$

dimana fungsi kumulatif *hazard* dirumuskan dengan persamaan 2.11.

$$H(t) = \int_0^t h(t) dt \quad (2.11)$$

Berdasarkan fungsi 2.10 dan 2.11 didapatkan rumus hubungan antara fungsi *survival* dan fungsi kumulatif *hazard* adalah sebagai berikut :

$$H(t) = -\ln S(t) \quad (2.12)$$

## 2.2 Kurva *Survival* Kaplan-Meier dan Uji *Log-Rank*

Kurva *survival* Kaplan-Meier merupakan analisis yang digunakan untuk menaksir fungsi *survival*. Metode ini didasarkan pada waktu kelangsungan hidup dan data sensor diasumsikan independen berdasarkan waktu kelangsungan hidupnya. Hal ini terjadi karena penyebab *failure time* tidak berhubungan dengan observasi yang disensor (Kleinbaum & Klein, 2005). Kurva ini menggambarkan sebuah hubungan antara estimasi fungsi *survival* pada waktu ke- $t$  dengan waktu *survival*. Estimasi fungsi *survival* dapat dibentuk menjadi persamaan berikut ini.

$$\hat{S}(t_{(j)}) = \hat{S}(t_{(j-1)}) \times \hat{Pr}[T > t_{(j)} | T \geq t_{(j)}] \quad (2.13)$$

Dimana :

$\hat{S}(t_{(j)})$  : kumulatif peluang untuk waktu *survival* pada waktu ke- $j$

$\hat{Pr}[T > t_{(j)} | T \geq t_{(j)}]$  : peluang *survival* untuk waktu lebih dari  $i$  jika peluang ini terjadi lebih dari atau saat waktu ke- $i$

Selanjutnya akan diuji dengan menggunakan uji *log-rank*. Pengujian ini merupakan analisis statistik non parametrik dimana digunakan ketika data tidak simetris atau data miring ke kanan. Uji *log-rank* dapat digunakan untuk melihat efisiensi perbandingan antara perawatan baru dengan perawatan lama apabila yang diukur adalah waktu hingga terjadi suatu peristiwa. Pengujian ini digunakan untuk membandingkan Kaplan-Meier dalam kelompok yang berbeda (Klein, 2012). Hipotesis yang digunakan untuk uji *log-rank* adalah sebagai berikut.

$H_0$  : Tidak ada perbedaan antar kelompok pada kurva *survival*.

$H_1$  : Adanya perbedaan antar kelompok pada kurva *survival*.

Statistik uji :

$$X^2_{hitung} = \sum_{g=1}^G \frac{(O_i - E_i)^2}{E_i} \quad (2.14)$$

dimana,

$$O_i - E_i = \sum_{f=1}^h (m_{if} - e_{if}) \quad (2.15)$$

$$e_{if} = \left( \frac{n_{if}}{\sum_{i=1}^G n_{if}} \right) \left( \sum_{i=1}^G m_{if} \right) \quad (2.16)$$

Keterangan :

$O_i$  : nilai observasi individu pada kelompok ke- $i$

$E_i$  : nilai ekspektasi individu pada kelompok ke- $i$

$m_{if}$  : jumlah objek yang mengalami *event* pada waktu ke- $t$  dan kelompok ke- $i$

- $n_{if}$  : jumlah objek yang bertahan pada waktu ke- $t$  dan kelompok ke- $i$   
 $e_{if}$  : nilai ekspektasi pada waktu ke- $t$  dan kelompok ke- $i$   
 $h$  : banyaknya observasi  
 $G$  : banyaknya kelompok  
 $f$  : *failure event* ( $1, 2, \dots, h$ )  
 $i$  :  $1, 2, \dots, G$

Kesimpulan: tolak  $H_0$  jika  $X^2_{hitung} > \chi^2_{(\alpha(G-1))}$

### 2.3 Regresi Cox

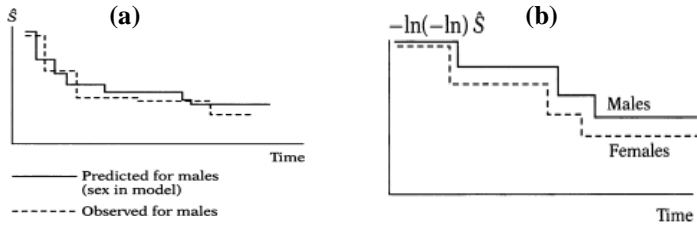
Metode ini merupakan analisis *survival* yang digunakan untuk mengetahui efek variabel independen terhadap variabel respon. Variabel respon dalam metode ini adalah waktu *survival* pada suatu objek terhadap suatu peristiwa tertentu (Cox, 1972). Regresi cox merupakan metode semiparametrik artinya modelnya terdapat komponen parametrik dan non parametrik. Metode ini baik digunakan jika distribusi dari waktu *survival* tidak diketahui pasti karena regresi ini tidak memiliki asumsi mengenai sifat dan bentuk sehingga hasil estimasi parameternya masih dapat dipercaya (Lee, 1980).

#### 2.3.1 Asumsi *Proportional Hazard*

Asumsi ini terpenuhi jika mempunyai nilai hazard ratio yang konstan dengan waktunya (Kleinbaum & Klein, 2005). Selain itu untuk mengetahui apakah asumsi ini terpenuhi atau tidak maka dapat digunakan beberapa pendekatan sebagai berikut :

##### a. Pendekatan Grafik

Grafik yang digunakan untuk penelitian *survival* adalah grafik  $\ln[-\ln S(t)]$  dan grafik *observed versus expected* kurva survival yang dijelaskan pada Gambar 2.4.



**Gambar 2.4** Ilustrasi Kurva Kaplan Meier

**Sumber :** Kleinbaum & Klein, 2005

Gambar 2.4 menunjukkan bahwa terdapat variabel jenis kelamin (*sex*). Selanjutnya akan dilakukan analisis apakah variabel *sex* memenuhi asumsi tersebut atau tidak. Pada Gambar 2.4 (a) asumsi *proportional hazard* terpenuhi karena kurva *survival* pengamatan dengan prediksi saling berimpit. Begitu juga dengan Gambar 2.4 (b) dapat dikatakan gambar tersebut memenuhi asumsi *proportional hazard* karena garis yang menunjukkan *males* sejajar dengan *females* (Kleinbaum & Klein, 2005).

b. Pendekatan Goodness of Fit

Pendekatan ini dilakukan dengan menggunakan pendekatan statistik uji dan *p-value* dengan hipotesis sebagai berikut (Kleinbaum & Klein, 2005).

$H_0$  : Asumsi *cox proportional hazard* terpenuhi.

$H_1$  : Asumsi *cox proportional hazard* tidak terpenuhi.

Dengan daerah penolakan yang berlaku yaitu :

i. *P-value* yang tidak signifikan adalah *p-value* yang bernilai lebih kecil dari 0,05 maka menunjukkan bahwa variabel yang diuji tolak  $H_0$

ii. Nilai  $\chi^2_{hitung} > \chi^2_{tabel}$  maka akan diambil kesimpulan tolak  $H_0$ .

### 2.3.2 Model Regresi Cox Proportional Hazard

Variabel independen dapat dilihat berpengaruh atau tidak terhadap pola waktu *survival* dengan menggunakan pemodelan regresi, pemodelan ini disebut dengan model regresi *cox*



*proportional hazard*. *Cox proportional hazard* atau regresi *cox* mempunyai hubungan antara variabel dependen dan independen dengan menggunakan fungsi *hazard* yang berbentuk eksponensial terhadap waktu tertentu. Pemodelan ini digunakan dengan tujuan untuk mengetahui efek variabel-variabel independen terhadap data *survival* secara bersamaan (Kleinbaum & Klein, 2005). Pemodelan tersebut dapat dituliskan sebagai berikut :

$$h(t, \mathbf{X}) = h_0(t) \exp \left[ \sum_{i=1}^p \beta_i x_i \right] \quad (2.17)$$

dengan :

$h_0(t)$  : *baseline hazard* pada waktu ke- $t$

$\beta_1, \beta_2, \dots, \beta_p$  : parameter regresi

$x_1, x_2, \dots, x_p$  : nilai dari variabel bebas

### 2.3.3 Estimasi Parameter Model Cox Proportional Hazard

Cara mengestimasi parameter dalam *pemodelan cox proportional hazard* adalah dengan memaksimumkan fungsi *partial likelihood*. Dalam menaksir model ini dibutuhkan penaksir koefisien variabel independen  $X_1, X_2, \dots, X_p$  dalam komponen linier model harus diketahui  $\beta_1, \beta_2, \dots, \beta_p$  dan fungsi *baseline hazard* yang akan ditaksir secara terpisah. Fungsi *baseline hazard* adalah analog dari istilah intercept pada regresi. Fungsi *baseline hazard* selalu bernilai positif. Apabila data terdiri atas  $n$  pengamatan waktu *survival* tak tersensor, ditunjukkan oleh  $t_{(1)}, t_{(2)}, \dots, t_{(n)}$  dan  $R(t_{(l)})$  adalah himpunan waktu yang beresiko pada waktu  $t_{(l)}$  yang terdiri dari semua obyek dengan waktu *survival* yang paling sedikit. Fungsi *likelihood*-nya adalah sebagai berikut.

$$L(\beta) = \prod_{i=1}^n \frac{\exp(\beta' \mathbf{x}_{(t)})}{\sum_{l \in R(t_{(l)})} \exp(\beta' \mathbf{x}_{(l)})} \quad (2.18)$$

dengan,

$\mathbf{x}_{(t)}$  : vektor variabel dari obyek yang gagal pada saat waktu ke- $t$ .

$R(t_{(l)})$  : himpunan obyek yang masih hidup pada waktu ke- $l$ .

$\mathbf{x}_{(l)}$  : vektor variabel obyek yang masih hidup dan merupakan elemen dari  $R(t_{(l)})$ .

Dari fungsi *likelihood* pada persamaan 2.18 akan dilanjutkan dengan membuat fungsi *ln-likelihood* seperti persamaan 2.19

$$\begin{aligned}
 \ln L(\boldsymbol{\beta}) &= \sum_{i=1}^n (\boldsymbol{\beta}' \mathbf{x}_{(i)}) - \sum_{i=1}^n \ln \left[ \sum_{l \in R(t_{(i)})} \exp(\boldsymbol{\beta}' \mathbf{x}_{(il)}) \right] \quad (2.19) \\
 &= \sum_{i=1}^n \sum_{l=1}^p (\beta_i x_{(il)}) - \sum_{i=1}^n \ln \left[ \sum_{l \in R(t_{(i)})} \exp(\beta_i x_{(il)}) \right] \\
 &= \sum_{i=1}^n \left( \sum_{l=1}^p (\beta_i x_{(il)}) - \sum_{l=1}^n \ln \left[ \sum_{l \in R(t_{(i)})} \exp(\beta_i x_{(il)}) \right] \right)
 \end{aligned}$$

Persamaan 2.19 dapat menghasilkan turunan pertama dan turunan kedua. Turunan pertama dari fungsi *likelihood* yang sudah ditransformasikan menjadi *ln-likelihood*. Turunan pertama dari fungsi *likelihood* dapat dilihat pada persamaan 2.20.

$$\begin{aligned}
 \frac{\partial \ln L(\beta_1, \beta_2, \dots, \beta_p)}{\partial \beta_1} &= \sum_{i=1}^n \left[ x_{1i} \frac{\sum_{l \in R(t_{(i)})} x_{1l} \exp \sum_{i=1}^p (\beta_i x_{(il)})}{\sum_{l \in R(t_{(i)})} \exp \sum_{i=1}^p (\beta_i x_{(il)})} \right] \quad (2.20) \\
 \frac{\partial \ln L(\beta_1, \beta_2, \dots, \beta_p)}{\partial \beta_2} &= \sum_{i=1}^n \left[ x_{2i} \frac{\sum_{l \in R(t_{(i)})} x_{2l} \exp \sum_{i=1}^p (\beta_i x_{(il)})}{\sum_{l \in R(t_{(i)})} \exp \sum_{i=1}^p (\beta_i x_{(il)})} \right] \\
 &\vdots \\
 \frac{\partial \ln L(\beta_1, \beta_2, \dots, \beta_p)}{\partial \beta_p} &= \sum_{i=1}^n \left[ x_{pi} \frac{\sum_{l \in R(t_{(i)})} x_{pl} \exp \sum_{i=1}^p (\beta_i x_{(il)})}{\sum_{l \in R(t_{(i)})} \exp \sum_{i=1}^p (\beta_i x_{(il)})} \right]
 \end{aligned}$$

Sedangkan turunan kedua dari dari fungsi *likelihood* yang sudah ditransformasikan menjadi *ln-likelihood* dapat dilihat pada persamaan 2.21.

$$\frac{\partial^2 \ln L(\boldsymbol{\beta})}{\partial \beta_j \partial \beta_{j^*}} = - \sum_{i=1}^n \left[ \frac{\sum_{l \in R(t_{(i)})} x_{jl} x_{j^*l} \exp \sum_{i=1}^p (\beta_i x_{(il)})}{\sum_{l \in R(t_{(i)})} \exp \sum_{i=1}^p (\beta_i x_{(il)})} \right] + \quad (2.21)$$

$$\sum_{i=1}^n \left[ \frac{\sum_{l \in R(t_{(i)})} x_{jl} \exp \sum_{i=1}^p (\beta_i x_{(il)}) \sum_{l \in R(t_{(i)})} x_{j^*l} \exp \sum_{i=1}^p (\beta_i x_{(il)})}{\sum_{l \in R(t_{(i)})} \exp \sum_{i=1}^p (\beta_i x_{(il)}) \sum_{l \in R(t_{(i)})} \exp \sum_{i=1}^p (\beta_i x_{(il)})} \right] \quad (2.21)$$

dengan  $j, j^* = 1, 2, \dots, p$

Estimasi parameter yang diperoleh dalam metode ini implisit sehingga digunakan metode iterasi numerik dengan Newton-Raphson (Collet, 2003). Jika  $g(\boldsymbol{\beta})$  merupakan vector yang berukuran  $p \times 1$  yang merupakan turunan pertama fungsi  $\ln L(\boldsymbol{\beta})$  terhadap parameter  $\boldsymbol{\beta}$ .  $H(\boldsymbol{\beta})$  adalah matrik hessian berukuran  $p \times p$  yang berisi turunan kedua dari fungsi  $\ln L(\boldsymbol{\beta})$ , maka estimasi parameter pada iterasi ke  $(l+1)$  terdapat pada persamaan 2.22.

$$\boldsymbol{\beta}^{(l+1)} = \boldsymbol{\beta}^{(l)} - H^{-1}(\boldsymbol{\beta}^{(l)}) g(\boldsymbol{\beta}^{(l)}) \quad (2.22)$$

Iterasi Newton-Rhapson akan berhenti jika  $\|\boldsymbol{\beta}^{(l+1)} - \boldsymbol{\beta}^{(l)}\| \leq \varepsilon$ , dimana  $\varepsilon$  merupakan suatu bilangan yang sangat kecil.

### 2.3.4 Seleksi Model Terbaik

Model terbaik didapatkan dengan cara seleksi sehingga dapat menggambarkan hubungan antara waktu *survival* dengan beberapa variabel independen secara tepat. Prosedur yang digunakan adalah metode eliminasi *backward* dengan menggunakan kriteria *Akaike Information Criteriation* (AIC). Model terbaik didapatkan dari kriteria nilai AIC terkecil dengan perhitungan pada persamaan 2.23.

$$AIC = 2k + 2L(\boldsymbol{\beta}) \quad (2.23)$$

Langkah-langkah dalam metode eliminasi *backward* dengan menggunakan kriteria AIC adalah sebagai berikut (Lee, 1980) :

1. Membuat model regresi yang berisi semua variabel prediktor yang tersedia.
2. Memilih satu variabel prediktor berdasarkan kriteria pemilihan.
3. Melakukan pengujian pada variabel prediktor.  
Mengulangi langkah 2 dan 3 sampai kriteria dalam pengujian tidak terpenuhi.

### 2.3.5 Pengujian Signifikansi Parameter

Pengujian ini dilakukan secara 2 tahap. Tahap pertama adalah pengujian serentak yang kemudian dilanjutkan oleh pengujian individual.

1. Uji Serentak

Hipotesis :

$$H_0 : \beta_1 = \beta_2 = \dots = \beta_p = 0$$

$$H_1 : \text{minimal ada satu } \beta_j \neq 0, j = 1, 2, \dots, p$$

Statistik uji :

$$G^2 = -2 \ln \frac{L(\hat{\omega})}{L(\hat{\Omega})} \quad (2.24)$$

$L(\hat{\omega})$  : nilai likelihood untuk model tanpa menyertakan variabel prediktor.

$L(\hat{\Omega})$  : nilai likelihood untuk model dengan menyertakan semua variabel prediktor.

$p$  : banyak parameter dalam model.

Keputusan : tolak  $H_0$  jika nilai statistik uji  $G^2 > \chi^2_{p, \alpha}$  atau

$$p\text{-value} < \alpha$$

2. Uji Parsial

Hipotesis :

$H_0 : \beta_j = 0$  (variabel prediktor tidak berpengaruh terhadap variabel respon)

$H_1 : \beta_j \neq 0$  ( variabel prediktor berpengaruh terhadap variabel respon),  $j = 1, 2, \dots, p$

Statistik Uji :

$$W^2 = \frac{(\hat{\beta}_j)^2}{(\text{SE}(\hat{\beta}_j))^2} \quad (2.25)$$

Tolak  $H_0$  jika  $W^2_{hit} > \chi^2_{\alpha, p}$  atau  $p\text{-value} < 0$ .

### 2.3.6 Hazard Ratio

*Hazard ratio* adalah tingkat resiko yang dilihat melalui perbandingan antara obyek dengan kondisi variabel prediktor pada kategori sukses atau gagal (Hosmer, Lemeshow, & May, 2008). Nilai eksponen dari koefisien regresi cox-nya merupakan estimasi dari *hazard rate*. Misal  $X$  adalah sebuah variabel prediktor dengan dua kategori, yaitu 0 dan 1. Hubungan antara variabel  $X$  dengan *hazard rate* atau  $h(t)$  dinyatakan dengan  $h_0(t|x) = h_0(t)e^\beta$ , maka Obyek dengan  $x = 1$ , fungsi hazard-nya:

$$h_0(t | x = 1) = h_0(t) e^{\beta(1)} = h_0(t) e^\beta$$

Obyek dengan  $x = 0$ , fungsi *hazard*-nya:

$$h_0(t | x = 0) = h_0(t) e^{\beta(0)} = h_0(t)$$

sehingga nilai *Hazard Ratio* dapat dihitung dengan rumus pada persamaan 2.25.

$$HR = \frac{h_0(t | x = 0)}{h_0(t | x = 1)} = \frac{h_0(t)e^\beta}{h_0(t)} = e^\beta \quad (2.26)$$

Nilai *hazard ratio* yang diperoleh tersebut memiliki arti bahwa tingkat kecepatan terjadinya *failure event* (laju kegagalan) pada obyek dengan kategori  $x=0$  adalah sebesar  $e^\beta$  kali tingkat kecepatan terjadinya resiko peristiwa *failure event* (laju kegagalan) pada obyek dengan kategori  $x=1$ .

### 2.3.7 Bootstrap

*Bootstrap* merupakan metode *resampling* nonparametrik yang bertujuan untuk menentukan estimasi standar *error* dan

*confidence interval* dari parameter populasi seperti *mean*, rasio, median, proporsi, koefisien korelasi atau koefisien regresi tanpa menggunakan asumsi distribusi. Metode *bootstrap* dapat digunakan untuk mengatasi data yang sedikit, data yang menyimpang dari asumsinya maupun data yang tidak memiliki asumsi distribusi (Sahinler & Topuz, 2007). Metode ini digunakan dengan mengambil sampel dari data asli dengan ukuran yang sama seperti data asli. Dalam metode *bootstrap* kedudukan data asli sama sebagai populasi. Teknik penyampelan seperti ini disebut *resampling bootstrap*.

*Resampling bootstrap* untuk estimasi rata-rata populasi dimisalkan sampel berukuran  $n$  yaitu  $X_1, X_2, X_3, \dots, X_n$  yang diambil dari populasi dan  $\hat{\mu}$  merupakan parameter  $\mu$  untuk dari data asli. Langkah-langkah dalam *resampling bootstrap* dapat dituliskan sebagai berikut.

1. Mengambil sampel *bootstrap* berukuran  $n$  secara acak dengan pengembalian sehingga data tersebut dapat terambil kembali sehingga dapat disebut sebagai sampel *bootstrap* pertama  $X^{*1}$ .
2. Menghitung  $\hat{\mu}_1^*$  dari statistik  $\hat{\mu}$  yang diinginkan dari sampel *bootstrap*  $X^{*1}$ .
3. Mengulangi langkah ke-1 dan ke-2 hingga  $B$  (replikasi yang diinginkan) kali dan akan diperoleh  $\hat{\mu}_1^*, \hat{\mu}_2^*, \dots, \hat{\mu}_B^*$ .
4. Pendekatan dalam estimasi *bootstrap* untuk  $\mu$  yaitu

$$\hat{\mu}^* = \sum_{b=1}^B \hat{\mu}_b^* \frac{1}{B}$$

Dalam mengestimasi *confidence interval bootstrap* untuk rata-rata diperlukan standar *error* yang dilambangkan dengan  $Se(\hat{\mu}^*)$ . Standar error tersebut akan membantu untuk mencari *confidence interval* dengan rumus sebagai berikut (Bennet, 2009).

$$\hat{\mu} + Z_{\alpha/2} SE_B(\hat{\mu}^*) < \hat{\mu} < \hat{\mu} + Z_{1-\alpha/2} SE_B(\hat{\mu}^*) \quad (2.27)$$

## **2.4 Kematian**

Kematian secara biologis merupakan berhentinya seluruh proses aktivitas dalam tubuh pada seorang individu. Hal ini ditandai dengan hilangnya fungsi otak, detak jantung berhenti, berhentinya aliran darah dan berhentinya proses pernafasan (Hartini, 2007). Kematian dapat dianggap sebagai peralihan kehidupan dari dunia menuju kehidupan lain, perspektif ini dapat dilihat dari sisi keagamaan yaitu agama Islam. Dalam agama Islam kematian didefinisikan sebagai kehilangan permanen dari fungsi integratif secara keseluruhan (Hasan, 2006). Kematian secara psikologis merupakan sebuah peristiwa yang sangat berpengaruh dalam kehidupan seseorang. Ada beberapa orang yang menganggap kematian merupakan sebuah malapetaka tetapi ada sebagian orang yang menganggap kehidupan di dunia hanya sementara karena ada kehidupan lain yang lebih mulia kelak yaitu di akhirat (Dister, 1982).

*“Halaman ini sengaja dikosongkan”*



## BAB III METODOLOGI PENELITIAN

### 3.1 Sumber Data

Sumber data yang digunakan pada penelitian ini adalah data sekunder yang didapatkan dari PT Taspen (Persero) Kantor Cabang Utama Surabaya. Data yang digunakan adalah data peserta pensiun meninggal aktif. Dalam penelitian ini, data yang digunakan berjumlah 100 data dari Februari 2017 sampai Maret 2018 dimana terdapat 60 pekerja yang meninggal. Pada penelitian ini data yang mengalami kematian adalah data yang memiliki kematian sebelum bulan Desember 2017.

### 3.2 Variabel Penelitian

Variabel penelitian yang digunakan dalam data penelitian peserta meninggal aktif ini terbagi menjadi dua macam yaitu variabel dependen dan variabel independen.

#### 1. Variabel Respon

Variabel respon dalam penelitian ini adalah data *survival* dari lama seseorang bekerja sampai meninggal. Hal ini akan dihitung dari waktu awal seseorang menjadi anggota Pegawai Negeri Sipil (*start point*) sampai meninggal (*end point*). Dalam penelitian ini data *survival* time akan di simbolkan dengan  $T$  dengan satuan waktunya adalah tahun.

**Tabel 3.1** Variabel Dependen Penelitian

Variabel	Deskripsi	Keterangan	Skala
$T$	Waktu Survival	Waktu dimana seseorang pekerja (PNS) bekerja sampai meninggal (tahun)	Rasio
$D$	Status Pekerja	1 : Data tidak tersensor ketika pekerja meninggal 0 : Data tersensor ketika pekerja belum meninggal.	Nominal

#### 2. Variabel Independen

Variabel independen atau yang biasa disebut covariate merupakan variabel yang diduga mempengaruhi faktor-faktor lama seseorang bekerja sampai meninggal, yaitu :

**Tabel 3.2** Variabel Independen Penelitian

Variabel	Deskripsi	Keterangan	Skala
$X_1$	Jenis Kelamin	0 : Perempuan 1 : Laki-laki	Nominal
$X_2$	Golongan Pekerja	0 : Golongan 1 & 2 3 : Golongan 3 4 : Golongan 4	Nominal
$X_3$	Pasangan	0 : Tidak memiliki istri / suami 1 : Jumlah istri/suami 1	Nominal
$X_4$	Anak	0 : Tidak memiliki anak 1 : Jumlah anak 1 2 : Jumlah anak 2	Nominal
$X_5$	Usia	Usia seorang peserta pegawai negeri sipil ketika meninggal (tahun).	Rasio

Struktur data yang digunakan dalam penelitian ini berada pada Tabel 3.3.

**Tabel 3.3** Struktur Data Penelitian

Peserta ke-	$T$	$D$	$X_1$	$X_2$	$X_3$	$X_4$	$X_5$
1	$T_1$	$D_1$	$X_{11}$	$X_{21}$	$X_{31}$	$X_{41}$	$X_{51}$
2	$T_2$	$D_2$	$X_{12}$	$X_{22}$	$X_{32}$	$X_{42}$	$X_{52}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
j	$T_j$	$D_j$	$X_{1j}$	$X_{2j}$	$X_{3j}$	$X_{4j}$	$X_{5j}$
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
n	$T_n$	$D_n$	$X_{1n}$	$X_{2n}$	$X_{3n}$	$X_{4n}$	$X_{5n}$

Keterangan :

$j$  : 1,2,3,...,n

$T_j$  : waktu survival pekerja ke-j

$D_j$  : status pekerja ke-j

$X_{1j}$  : jenis kelamin pekerja ke-j

$X_{2j}$  : golongan pekerja ke-j

$X_{3j}$  : jumlah pasangan pekerja ke- $j$

$X_{4j}$  : jumlah anak pekerja ke- $j$

$X_{5j}$  : usia pekerja ke- $j$

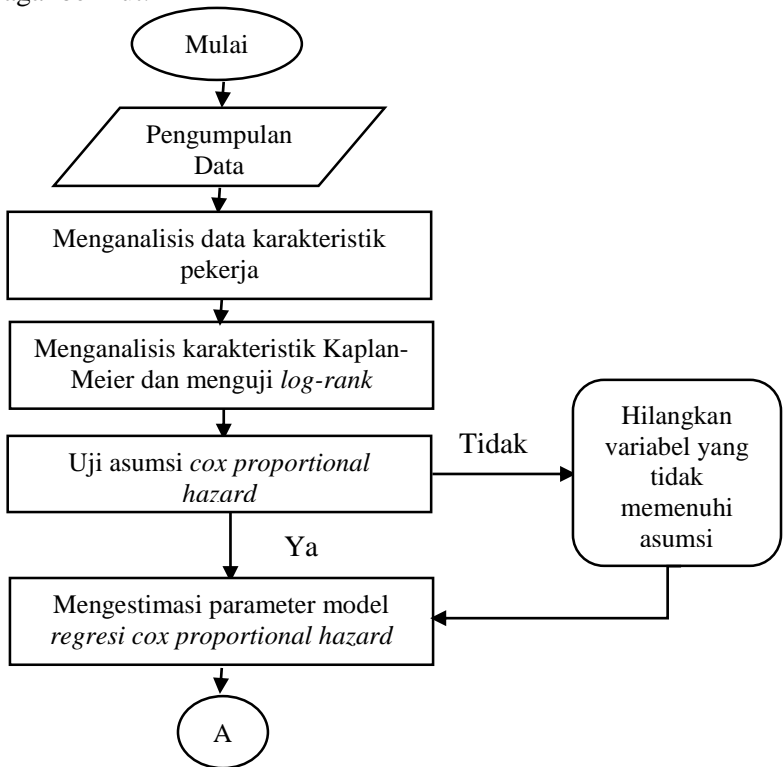
### 3.3 Langkah-langkah Analisis Data

Langkah analisis yang akan dilakukan dalam penelitian ini adalah sebagai berikut :

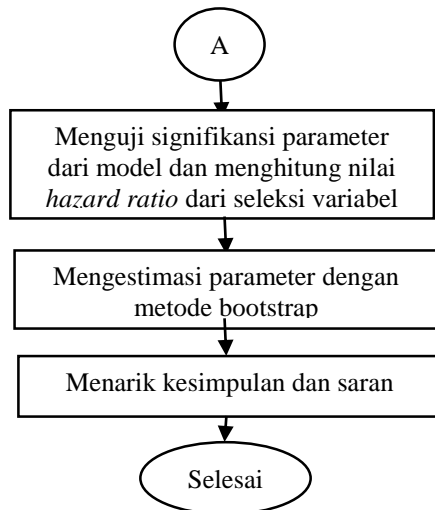
1. Mengumpulkan data dari PT. Taspen (Persero) Kantor Cabang Utama Surabaya.
2. Melakukan analisis statistika deskriptif untuk mengetahui karakteristik data pekerja. Langkah analisis yang akan dilakukan adalah sebagai berikut.
  - a. Menghitung nilai *mean*, standar deviasi, nilai maksimum dan nilai minimum pada variabel dependen ( $T$ ) dan variabel independen yaitu usia ( $X_5$ ).
  - b. Menghitung nilai presentase dengan membuat *pie chart* pada variabel independen jenis kelamin ( $X_1$ ), Golongan Pekerja ( $X_2$ ), Pasangan ( $X_3$ ), dan Anak ( $X_4$ ). Hal ini dilakukan untuk menganalisis karakteristik variabel independen.
3. Menganalisis karakteristik kurva Kaplan-Meier dan melakukan pengujian *Log-Rank* untuk semua faktor-faktor yang mempengaruhi waktu *survival* pekerja.
4. Membuat model regresi *Cox Proportional Hazard* dengan menggunakan langkah-langkah berikut.
  - a. Pemeriksaan asumsi *Proportional Hazard* terhadap semua variabel independen dengan menggunakan *software R*.
  - b. Melakukan estimasi parameter model regresi *Cox Proportional Hazard*.
  - c. Melakukan pengujian signifikansi parameter.
  - d. Melakukan seleksi model terbaik dengan menggunakan pengujian AIC.
  - e. Model yang dibentuk berdasarkan nilai AIC yang terkecil.

- f. Melakukan penghitungan nilai *Hazard Ratio* variabel untuk menginterpretasikan model
  - g. Melakukan proses *bootstrap* sebanyak bilangan  $B$  kali.
  - h. Membuat *confidence interval* untuk menentukan variabel tersebut signifikan atau tidak.
5. Membuat kesimpulan dan saran dari hasil yang sudah didapatkan.

Langkah-langkah analisis ini dapat diubah menjadi diagram alir sebagai berikut.



**Gambar 3.1** Diagram Alir



**Gambar 3.2** Diagram Alir Lanjutan

*“Halaman ini sengaja dikosongkan”*

## BAB IV ANALISIS DAN PEMBAHASAN

Pada bab ini akan dijelaskan bagaimana hasil pengolahan metode-metode yang telah dijelaskan pada tinjauan pustaka. Hasil pengolahan tersebut akan diinterpretasikan berdasarkan hasil *output software* R dan SPSS.

### 4.1 Analisis Karakteristik Variabel pada Data Pekerja

Analisis karakteristik variabel dilakukan dengan menggunakan statistika deskriptif. Analisis ini bertujuan untuk mengetahui karakteristik dari variabel dependen dan variabel independen yang dipakai dalam penelitian ini.

#### 4.1.1 Karakteristik Waktu *Survival* ( $T$ ) dan Usia ( $X_5$ )

Pengolahan statistika deskriptif terhadap variabel waktu *survival* dan variabel usia dilakukan dengan menggunakan *software* Minitab sehingga dihasilkan *output* statistika deskriptif.

**Tabel 4.1** Statistika Deskriptif untuk Data Pekerja

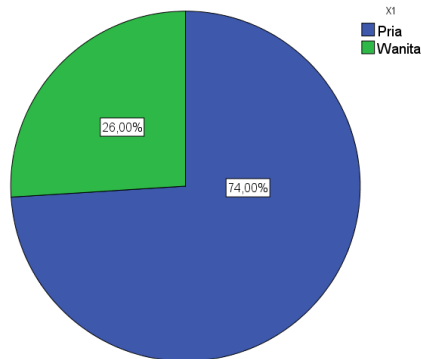
Variabel	Rata-Rata	Standar Deviasi	Minimal	Maksimal
Waktu <i>Survival</i>	24,575	9,753	3,408	39,975
Usia	52,575	5,927	30,088	62,532

Tabel 4.1 menunjukkan bahwa nilai rata-rata masa kerja untuk pekerja (waktu *survival*) adalah 24 tahun dengan standar deviasi sebesar 10 tahun. Masa kerja paling sedikit adalah 3 tahun sedangkan masa kerja paling lama sebesar 40 tahun. Hal ini dapat menginformasikan bahwa ketahanan seorang pekerja dalam bekerja sampai meninggal yaitu 40 tahun.

Variabel lainnya yang dianalisis dalam sub bab ini adalah usia. Rata-rata usia seorang pekerja ketika meninggal adalah 53 tahun dengan standar deviasi sebesar 6 tahun. Tetapi dalam penelitian ini terdapat seorang pekerja meninggal ketika berumur 30 tahun. Usia terlama seorang pekerja meninggal ketika berumur 63 tahun.

#### 4.1.2 Karakteristik Jenis Kelamin ( $X_1$ )

Kematian seorang pekerja dapat terjadi karena beberapa faktor. Salah satu faktor yang diduga mempengaruhi kematian pekerja adalah perbedaan jenis kelamin. Karakteristik variabel jenis kelamin akan diperlihatkan melalui diagram *pie*.



**Gambar 4.1** *Pie Chart* Jenis Kelamin

Gambar 4.1 menunjukkan bahwa pada 100 data pekerja pada periode bulan Februari 2017 sampai Maret 2018 dengan jenis kelamin pria lebih banyak daripada jenis kelamin wanita. Hal ini dibuktikan dengan area berwarna biru lebih luas daripada area berwarna hijau. Persentase pekerja wanita dari data ini sebesar 26% artinya dari 100 orang pekerja terdapat 26 orang pekerja berjenis kelamin wanita. Sedangkan pekerja dengan jenis kelamin pria memiliki persentase 74%.

**Tabel 4.2.** Tabulasi Silang Jenis Kelamin dengan Status Pekerja

Status Pekerja	Jenis Kelamin		Total
	Perempuan	Pria	
Tersensor	7	33	40
Tidak Tersensor	19	30	60
<b>Total</b>	26	74	100

Tabel 4.2 menunjukkan bahwa hasil tabulasi silang antara jenis kelamin dan status pekerja adalah dari 26 pekerja perempuan terdapat 7 orang pekerja yang masih hidup dan 19 orang pekerja yang telah meninggal. Pekerja dengan jenis kelamin laki-laki yang



sudah meninggal sebesar 30 orang dari total 74 pria yang diteliti. Sedangkan pekerja dengan jenis kelamin pria yang belum meninggal sebesar 33 orang.

**Tabel 4.3 Tabulasi Silang Rata-Rata Waktu *Survival* dan Usia pada Variabel Jenis Kelamin**

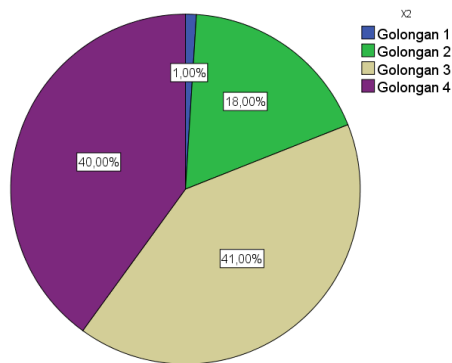
Variabel	Status Pekerja	Jenis Kelamin	
		Perempuan	Pria
<b>Waktu <i>Survival</i></b>	<b>Tersensor</b>	19,00	27,34
	<b>Tidak Tersensor</b>	24,37	23,39
	<b>Semua</b>	22,92	25,15
<b>Usia</b>	<b>Tersensor</b>	49,97	54,464
	<b>Tidak Tersensor</b>	51,84	51,84
	<b>Semua</b>	51,34	53,008

Jika variabel jenis kelamin dikaitkan dengan waktu *survival*-nya dapat dilihat pada Tabel 4.3 maka rata-rata pekerja dengan jenis kelamin pria memiliki waktu ketahanan hidup sebesar 23 tahun sedangkan pekerja dengan jenis kelamin wanita memiliki waktu ketahanan hidup sebesar 25 tahun. Namun jika diolah menurut status pekerjaanya rata-rata masa kerja wanita yang masih hidup dalam data 100 orang pekerja adalah 19 tahun sedangkan yang meninggal sebesar 24 tahun. Jika dilihat dari status pekerja pada jenis kelamin pria, waktu ketahanan hidup pekerja yang belum meninggal dan yang sudah meninggal berturut-turut sebesar 27 tahun dan 23 tahun.

Selanjutnya variabel jenis kelamin akan dikaitkan dengan usia maka pekerja jenis kelamin wanita cenderung memiliki rata-rata usia lebih lama daripada pria. Rata-rata usia pekerja wanita sebesar 53 tahun dan rata-rata usia pria sebesar 51 tahun. Jika dilihat dari status pekerja pada jenis kelamin wanita, data tersensor pada usia wanita lebih kecil daripada data tidak tersensor. Berbeda dengan jenis kelamin pria, pada status pekerja data tersensor usianya lebih lama daripada usia data tidak tersensor.

#### 4.1.3 Karakteristik Golongan Pekerja ( $X_2$ )

Golongan pekerjaan diduga menjadi faktor penyebab kematian dikarenakan semakin besar tuntutan pekerjaan yang didapat maka akan semakin berat beban pekerjaannya. Dalam penelitian ini golongan pekerja akan dibagi menjadi 4 bagian. Golongan 1 untuk Pegawai Negeri Sipil (PNS) dengan golongan 1a, 1b, 1c, dan 1d. Golongan 2 untuk PNS dengan golongan 2a, 2b, 2c, dan 2d. Golongan 3 untuk PNS dengan golongan 3a, 3b, 3c dan 3d. Sedangkan golongan 4 untuk PNS dengan golongan 4a, 4b, 4c dan 4d.



**Gambar 4.2** *Pie Chart* Golongan Pekerja

Berdasarkan Gambar 4.2 dari 100 data, golongan yang paling banyak dijabat oleh pekerja adalah golongan 3 sebesar 41 orang pekerja dengan persentase 41%. Kemudian dilanjutkan dengan golongan 4 yang hanya memiliki selisih satu pekerja. Peringkat selanjutnya adalah pekerja dengan golongan 2 dimana golongan ini memiliki persentase sebesar 18% dengan jumlah 18 orang pekerja. Persentase terkecil dari variabel golongan pekerja adalah pekerja dengan golongan 1 yang hanya memiliki persentase 1% setara dengan 1 orang pekerja. Dalam kasus ini, karena pekerja golongan 1 hanya memiliki 1 pekerja maka akan digabungkan dengan pekerja golongan 2.

**Tabel 4.4.** Tabulasi Silang Status Pekerja dengan Golongan

Status Pekerja	Golongan			Total
	Golongan 1&2	Golongan 3	Golongan 4	
Tersensor	8	11	21	40
Tidak Tersensor	11	30	19	60
<b>Total</b>	19	41	40	100

Berdasarkan Tabel 4.4 terlihat bahwa jumlah pekerja terbanyak yang memiliki kriteria data tersensor adalah golongan 4 dengan jumlah sebesar 21 namun jumlah pekerja tersedikit dengan kriteria data tersensor adalah pekerja dengan golongan 1 dan 2. Kriteria data tidak tersensor atau pekerja yang sudah meninggal yang paling banyak terlihat pada variabel kategorik golongan 3 dengan nilai sebesar 30, sebaliknya jumlah pekerja pada golongan 1 dan 2 memiliki jumlah pekerja paling sedikit untuk data pekerja yang sudah meninggal.

**Tabel 4.5.** Tabulasi Silang Rata-Rata Waktu *Survival* dan Usia pada Variabel Golongan Pekerja

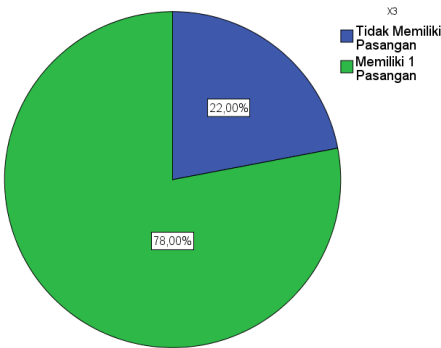
Variabel	Status Pekerja	Golongan		
		Golongan 1&2	Golongan 3	Golongan 4
<b>Waktu Survival</b>	Tersensor	10,680	25,67	31,78
	Tidak Tersensor	18,58	21,42	30,27
	Semua	15,26	22,56	31,067
<b>Usia</b>	Tersensor	48,05	54,45	55,417
	Tidak Tersensor	46,37	51,579	55,413
	Semua	47,08	52,35	55,415

Jika variabel golongan dikaitkan dengan waktu *survival*-nya dapat dilihat pada Tabel 4.5 rata-rata waktu ketahanan pekerja dari 100 data pada data tersensor yang paling lama terdapat pada golongan 4. Hal ini berlaku pada data tidak tersensor atau pekerja yang sudah meninggal. Sedangkan rata-rata waktu ketahanan seorang pekerja yang paling pendek pada data tersensor maupun tidak tersensor adalah pekerja dengan golongan 1 dan 2.

Selanjutnya variabel golongan akan dikaitkan dengan usia maka pekerja dengan status pekerja data tersensor yang memiliki usia terpanjang adalah pekerja dengan golongan 4, hal ini berlaku juga untuk data tidak tersensor. Sedangkan usia terpendek pada data tersensor dan tidak tersensor dari variabel golongan berada pada pekerja dengan golongan 1 dan 2.

4.1.4 Karakteristik Pasangan ( $X_3$ )

Variabel pasangan diduga dapat menjadi pemicu kematian pekerja. Karakteristik variabel pasangan akan ditampilkan dengan diagram *pie*.



Gambar 4.3 *Pie Chart* Pasangan

Kategori yang paling dominan pada Gambar 4.3 adalah pekerja yang memiliki 1 pasangan. Pekerja yang mempunyai 1 pasangan memiliki persentase sebesar 78% setara dengan 78 orang pekerja. Sedangkan pekerja yang tidak memiliki pasangan memiliki persentase sebesar 22% setara dengan 22 orang. Hal ini menunjukkan bahwa kebanyakan pekerja memiliki 1 pasangan ketika meninggal.

Tabel 4.6. Tabulasi Silang Status Pekerja dengan Pasangan

Status Pekerja	Pasangan		
	Tidak Punya Pasangan	Tidak Punya Pasangan	Pasangan
Tersensor	6	34	40
Tidak Tersensor	16	44	60
Total	22	78	100

Tabel 4.6 menunjukkan bahwa jumlah data tersensor terbanyak terdapat pada variabel pasangan dengan kategori pekerja yang memiliki 1 pasangan dengan jumlah sebesar 34 orang. Data tidak tersensor dengan jumlah pekerja paling banyak terdapat pada variabel pasangan dengan kategori punya 1 pasangan dengan jumlah 44 orang pekerja.

**Tabel 4.7.** Tabulasi Silang Rata-Rata Waktu *Survival* dan Usia pada Variabel Pasangan

Variabel	Status Pekerja	Pasangan	
		Tidak Punya Pasangan	Punya 1 Pasangan
Waktu <i>Survival</i>	Tersensor	21,35	26,68
	Tidak Tersensor	19,45	25,25
	Semua	19,97	25,87
Usia	Tersensor	52,23	53,933
	Tidak Tersensor	50,34	52,385
	Semua	50,85	53,060

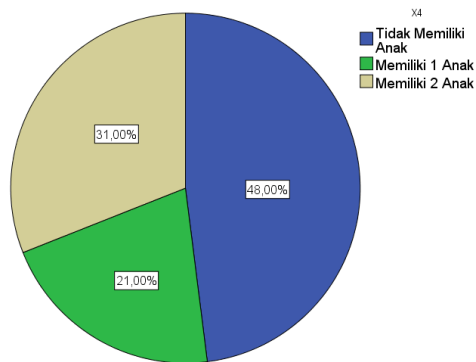
Tabel 4.7 menunjukkan bahwa variabel pasangan jika dikaitkan dengan waktu *survival* didapatkan hasil bahwa pekerja yang tidak memiliki pasangan memiliki rata-rata waktu ketahanan 20 tahun. sedangkan untuk pekerja yang memiliki satu pasangan mempunyai waktu ketahanan selama 26 tahun. Namun jika dikelompokkan berdasarkan status pekerja, waktu ketahanan hidup pekerja untuk data tersensor yang memiliki satu pasangan bernilai 27 tahun sedangkan yang tidak memiliki pasangan bernilai sebesar 21 tahun. Sedangkan untuk pekerja dengan status pekerja data tidak tersensor memiliki waktu ketahanan selama 25 tahun untuk pekerja yang memiliki satu pasangan dan 19 tahun untuk pekerja yang tidak memiliki pasangan.

Jika variabel pasangan dikaitkan dengan variabel usia didapatkan hasil bahwa pekerja yang tidak memiliki pasangan memiliki rata-rata usia meninggal 51 tahun sedangkan rata-rata untuk pekerja yang memiliki pasangan memiliki usia kematian sebesar 53 tahun. Usia pekerja dapat dikelompokkan berdasarkan status pekerja, data tersensor dengan variabel pasangan pada kategori tidak punya pasangan memiliki rata-rata usia sebesar 52

tahun dan kategori pekerja yang memiliki satu pasangan sebesar 54 tahun. Berdasarkan data tidak tersensor variabel pasangan yang memiliki kategori tidak mempunyai pasangan memiliki rata-rata usia 50 tahun dan untuk kategori pekerja yang memiliki satu pasangan sebesar 53 tahun.

#### 4.1.5 Karakteristik Anak ( $X_4$ )

Anak merupakan sebuah anugerah yang diberikan kepada pasangan suami istri. Namun dengan perkembangan ekonomi yang tidak menentu diduga dapat memicu kematian orang tua. Karakteristik variabel anak akan ditampilkan dengan menggunakan *pie chart*.



**Gambar 4.4** *Pie Chart Anak*

Gambar 4.4 menunjukkan bahwa pekerja yang tidak memiliki anak menempati peringkat pertama berlawanan dengan dugaan awal. Hal ini menunjukkan pekerja yang tidak memiliki anak pada waktu penelitian ini tingkat kematiannya lebih tinggi artinya pekerja yang tidak memiliki anak memiliki alasan lain untuk meninggal. Namun dalam penelitian ini hanya akan menganalisis apakah variabel anak akan berpengaruh pada faktor kematian pekerja pada bulan ini. Persentase untuk jumlah pekerja yang tidak memiliki anak sebesar 48% yang setara dengan 48 orang. Sedangkan peringkat berikutnya disusul oleh pekerja yang

memiliki 2 anak dengan persentase 31% setara dengan 31 orang. Pekerja yang memiliki 1 anak memiliki persentase terkecil dengan jumlah 21 orang atau setara dengan persentase sebesar 21%.

**Tabel 4.8.** Tabulasi Silang Status Pekerja dengan Anak

Status Pekerja	Anak			Total
	Tidak Memiliki Anak	Memiliki 1 Anak	Memiliki 2 Anak	
<b>Tersensor</b>	22	8	10	40
<b>Tidak Tersensor</b>	26	13	21	60
<b>Total</b>	48	21	31	100

Tabel 4.8 menunjukkan bahwa jumlah pekerja terbanyak pada data tersensor berada pada variabel anak dengan kategori yang tidak memiliki anak dan jumlah pekerja terkecil pada data tersensor adalah variabel anak dengan kategori pekerja yang memiliki 1 anak. Data tidak tersensor pada variabel anak dengan jumlah pekerja terbanyak terdapat pada kategori pekerja yang tidak memiliki anak dan jumlah pekerja tersedikit terdapat pada kategori pekerja yang memiliki satu anak.

**Tabel 4.9.** Tabulasi Silang Rata-Rata Waktu *Survival* dan Usia pada Variabel Anak

Variabel	Status Pekerja	Anak		
		Tidak Memiliki Anak	Memiliki 1 Anak	Memiliki 2 Anak
<b>Waktu Survival</b>	<b>Tersensor</b>	27,55	23,15	24,41
	<b>Tidak Tersensor</b>	24,77	24,52	21,88
	<b>Semua</b>	26,04	24,00	22,70
<b>Usia</b>	<b>Tersensor</b>	54,78	51,59	52,93
	<b>Tidak Tersensor</b>	53,04	51,61	50,49
	<b>Semua</b>	53,839	51,60	51,278

Tabel 4.9 menunjukkan bahwa variabel anak jika dikaitkan dengan waktu *survival* didapatkan hasil bahwa pekerja yang tidak mempunyai anak memiliki rata-rata waktu ketahanan 26 tahun, pekerja yang memiliki 1 anak mempunyai waktu ketahanan selama 24 tahun dan untuk pekerja yang memiliki dua anak mempunyai

waktu ketahanan selama 23 tahun. Namun jika dikelompokkan berdasarkan status pekerja, waktu ketahanan hidup pekerja untuk data tersensor yang paling lama adalah pekerja yang tidak memiliki anak. Hal tersebut berlaku untuk data tidak tersensor dari 100 data pekerja.

Jika variabel anak dikaitkan dengan variabel usia didapatkan hasil bahwa pekerja berdasarkan status pekerja untuk data tersensor yang memiliki rata-rata usia terlama dari 100 data pekerja adalah variabel anak dengan kategori yang tidak memiliki anak. Sedangkan untuk data tidak tersensor rata-rata usia terlama juga berada pada variabel anak dengan kategori yang tidak memiliki anak.

## **4.2 Kurva Kaplan-Meier dan Pengujian *Log-Rank* untuk Semua Variabel Independen yang Mempengaruhi *Survival Time***

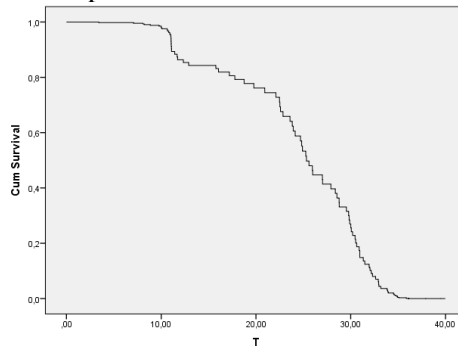
Analisis karakteristik kurva Kaplan-Meier ini digunakan untuk mengetahui karakteristik kurva *survival* pekerja yang sudah meninggal berdasarkan faktor-faktor yang diduga mempengaruhi ketahanan hidup pekerja tersebut, sedangkan pengujian *Log-Rank* digunakan untuk membedakan grup yang berbeda dalam satu kurva *survival*. Analisis *survival* memiliki dua cara untuk menganalisis penelitian ini yaitu dengan pengamatan tersensor dan tidak tersensor. Pengamatan tidak tersensor adalah pengamatan dimana dalam sebuah data individu yang diambil hanya yang meninggal saja atau yang mengalami kejadian yang diamati. Sering terjadi dalam suatu data waktu *survival* ketika di analisis tidak memenuhi asumsi normalitas dan tidak diketahui distribusi populasinya sehingga diperlukan analisis dengan metode non parametrik (Lee & Wang, 2003). Berikut ini akan diolah 100 data pekerja yang sudah meninggal dengan kurva Kaplan-Meier dan uji *Log-Rank*.

### **4.2.1 Karakteristik Pekerja Berdasarkan Semua Variabel Independen**

Sebelum mengetahui karakteristik dan perbedaan kurva *survival* pekerja berdasarkan faktor-faktor yang mempengaruhinya,



terlebih dahulu dilakukan analisis dengan kurva Kaplan-Meier secara umum. Hal ini dilakukan untuk mengetahui gambaran karakteristik kurva secara umum dari semua variabel independen. Berikut ini kurva Kaplan-Meier dari semua variabel independen

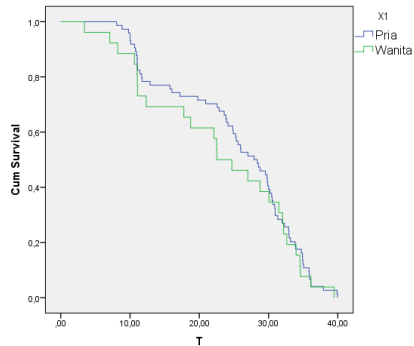


**Gambar 4.5** Kurva Kaplan-Meier Semua Variabel Independen

Gambar 4.5 menunjukkan bahwa kurva *survival* rata-rata dari semua faktor-faktor yang diduga mempengaruhi waktu *survival* pekerja turun sampai pada rentan waktu 40 tahun. Hal ini menunjukkan bahwa rentang waktu 100 orang pekerja mengalami kematian adalah 40 tahun. Pada rentang waktu 0 sampai sebelum 10 tahun, pekerja memiliki probabilitas untuk bertahan hidup sebesar 1. Namun dengan semakin lamanya waktu bekerja maka probabilitas kemungkinan pekerja untuk bertahan hidup semakin menurun.

#### 4.2.2 Karakteristik Pekerja Berdasarkan Variabel Kategorik Jenis Kelamin ( $X_1$ ) dan Uji *Log-Rank*

Jenis kelamin merupakan variabel kategorik yang terdiri atas pria dan wanita. Kurva Kaplan-Meier digunakan untuk mengetahui tingkat kematian pekerja untuk setiap kategori jenis kelamin. Selain itu kurva Kaplan-Meier digunakan untuk menduga apakah setiap kategori memiliki perbedaan kurva *survival*-nya. Berikut ini akan ditampilkan pada Gambar 4.6 kurva Kaplan-Meier untuk variabel kategorik jenis kelamin yang diolah dengan menggunakan *software* SPSS.



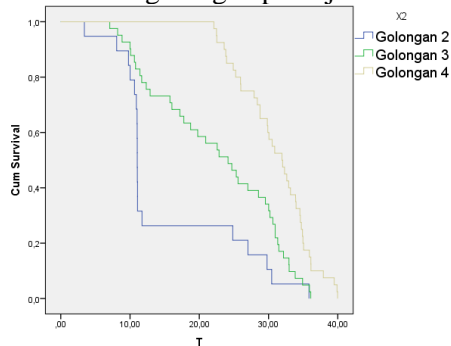
**Gambar 4.6** Kurva Kaplan-Meier untuk Variabel Jenis Kelamin

Berdasarkan Gambar 4.6 ditunjukkan bahwa garis yang berwarna biru merupakan pekerja dengan berjenis kelamin pria, sedangkan garis berwarna hijau merupakan pekerja dengan berjenis kelamin wanita. Dari gambar tersebut terlihat bahwa garis biru berada di atas garis hijau. Hal ini menunjukkan bahwa tingkat kematian pekerja berjenis kelamin wanita lebih tinggi daripada pria. Fungsi *survival* antar kelompok pekerja berjenis kelamin pria maupun wanita yang terlihat pada Gambar 4.6 kurva saling berimpit atau berdekatan, artinya diduga tidak ada perbedaan pada kurva *survival* yang memiliki jenis kelamin pria maupun wanita.

Pendugaan ini akan diperkuat dengan pengujian *Log-Rank*. Uji *Log-Rank* pada analisis ini akan menggunakan *p-value* dimana jika *p-value* lebih besar dari nilai  $\alpha$  sebesar 0,05 maka akan menghasilkan keputusan gagal tolak  $H_0$ . *P-value* untuk variabel jenis kelamin sebesar 0,541 lebih besar dari nilai  $\alpha$  artinya diambil keputusan gagal tolak  $H_0$ . Selain itu, pengujian *log-rank* dapat menggunakan nilai *chi-square* yang akan dibandingkan dengan nilai *chi-square* tabel. Nilai *chi-square* pada pengujian ini sebesar 0,374 lebih kecil dari nilai *chi-square* tabel yang bernilai 3,841, artinya keputusan yang akan diambil gagal tolak  $H_0$ . Maka dapat disimpulkan tidak ada perbedaan antara kurva *survival* pekerja yang berjenis kelamin pria maupun wanita.

#### 4.2.3 Karakteristik Pekerja Berdasarkan Variabel Kategorik Golongan Pekerja ( $X_2$ ) dan Uji *Log-Rank*

Golongan pekerja dalam variabel ini terbagi menjadi empat macam. Berikut ini akan dianalisis *output software* SPSS yaitu kurva Kaplan-Meier dari golongan pekerja.



**Gambar 4.7** Kurva Kaplan-Meier untuk Golongan Pekerja

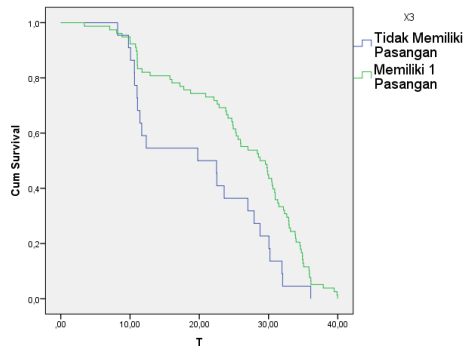
Berdasarkan Gambar 4.7 ditunjukkan bahwa garis yang berwarna biru merupakan pekerja dengan golongan 2, garis berwarna hijau merupakan pekerja dengan golongan 3, dan garis berwarna kuning merupakan pekerja dengan golongan 4. Dari gambar tersebut terlihat bahwa garis kuning berada di atas garis lainnya, hal ini menunjukkan bahwa pekerja dengan golongan 4 cenderung bisa bertahan hidup lebih lama dibandingkan dengan kedua kategori lainnya. Fungsi *survival* antar kelompok pekerja yang terlihat pada Gambar 4.7 kurva tidak saling berimpit atau berpotongan, artinya diduga ada perbedaan kurva *survival* antara golongan pekerja 1 & 2, 3, dan 4.

Pendugaan ini akan diperkuat dengan pengujian *Log-Rank*. Pengujian *Log-Rank* pada variabel golongan pekerja akan menggunakan *p-value*. Jika keputusan yang diambil pada pengujian ini adalah gagal tolak  $H_0$  maka nilai *p-value* lebih besar dari nilai  $\alpha$  sebesar 0,05. *P-value* dari hasil *software* SPSS sebesar 0,000 lebih kecil dari nilai  $\alpha$  sehingga akan diambil keputusan tolak  $H_0$  artinya ada perbedaan pada kurva *survival* dengan golongan pekerja 2, 3 dan 4. Selain itu, pengujian *log-rank* dapat

menggunakan nilai *chi-square* yang akan dibandingkan dengan nilai *chi-square* tabel. Nilai *chi-square* pada pengujian ini sebesar 33,428 lebih besar dari nilai *chi-square* tabel yang bernilai 3,841, artinya keputusan yang akan diambil tolak  $H_0$ . Maka dapat disimpulkan kurva *survival* pekerja memiliki perbedaan secara signifikan berdasarkan faktor golongan pekerja.

#### 4.2.4 Karakteristik Pekerja Berdasarkan Variabel Kategorik Pasangan ( $X_3$ ) dan Uji Log-Rank

Pasangan merupakan variabel kategorik yang terdiri dari pekerja yang meninggalkan pasangan ketika meninggal dan pekerja yang memiliki 1 pasangan ketika meninggal. Berikut ini akan ditampilkan kurva Kaplan-Meier untuk variabel kategorik pasangan.



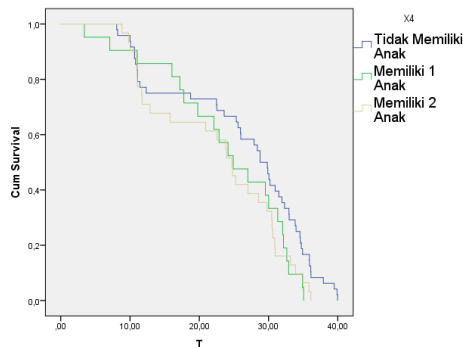
**Gambar 4.8** Kurva Kaplan-Meier untuk Variabel Pasangan

Informasi yang bisa didapatkan dari Gambar 4.8 adalah garis yang berwarna biru merupakan pekerja yang tidak memiliki pasangan, sedangkan garis berwarna hijau merupakan pekerja yang memiliki satu pasangan. Pada kurva tersebut terlihat bahwa garis hijau berada di atas garis biru, artinya tingkat kematian pekerja yang tidak memiliki pasangan lebih tinggi daripada pekerja yang memiliki satu pasangan. Fungsi *survival* antar kelompok pekerja yang tidak memiliki pasangan dan memiliki pasangan pada Gambar 4.8 kurva tidak saling berimpit atau berpotongan, artinya diduga ada perbedaan kurva *survival* antara pasangan yang tidak memiliki pasangan maupun yang memiliki satu pasangan.

Hasil pendugaan ini akan dibuktikan dengan uji *log-rank*. Pembuktian dari dugaan kurva Kaplan-Meier akan diuji dengan menggunakan *p-value* yang didapatkan dari *software* SPSS dimana jika *p-value* lebih besar dari nilai taraf signifikans 5% maka akan menghasilkan keputusan gagal tolak  $H_0$ . *P-value* dari variabel ini adalah 0,004 lebih kecil dari taraf signifikan. Selain itu, pengujian *log-rank* dapat menggunakan nilai *chi-square* yang akan dibandingkan dengan nilai *chi-square* tabel. Nilai *chi-square* pada pengujian ini sebesar 8,077 lebih besar dari nilai *chi-square* tabel yang bernilai 3,841, artinya keputusan yang akan diambil tolak  $H_0$ . Maka dapat disimpulkan ada perbedaan kurva *survival* antara pekerja yang tidak memiliki pasangan dan yang memiliki satu pasangan.

#### 4.2.5 Karakteristik Pekerja Berdasarkan Variabel Kategorik Anak ( $X_4$ ) dan Uji *Log-Rank*

Anak merupakan variabel kategorik yang terdiri dari tiga level dimana level pertama adalah ketika pekerja tidak memiliki anak, level kedua adalah ketika pekerja memiliki anak 1, dan pekerja ketiga adalah ketika pekerja memiliki anak 2. Berikut ini akan ditampilkan kurva Kaplan-Meier untuk variabel kategorik anak.



**Gambar 4.9** Kurva Kaplan-Meier untuk Variabel Anak

Gambar 4.9 menunjukkan bahwa garis yang berwarna biru merupakan pekerja yang tidak memiliki anak, ketika pekerja

memiliki satu anak maka garis berwarna hijau, dan yang terakhir garis berwarna kuning merupakan pekerja yang memiliki dua anak. Garis berwarna kuning atau pekerja yang memiliki dua anak berada di bawah garis lainnya. Hal ini menunjukkan bahwa tingkat kematian pekerja yang memiliki dua anak lebih tinggi daripada pekerja yang tidak memiliki anak dan memiliki 1 anak. Kurva-kurva pada Gambar 4.9 terlihat saling berimpit atau berdekatan, diduga tidak ada perbedaan pada kurva *survival* antar kelompok variabel anak.

Kurva Kaplan-Meier pada Gambar 4.9 akan dibuktikan dengan *p-value* dari uji *log-rank* pada hasil perhitungan *software* SPSS. *P-value* diketahui sebesar 0,061. Jika *p-value* dibandingkan dengan taraf signifikan 5% maka akan dihasilkan keputusan gagal tolak  $H_0$ . Selain itu, pengujian *log-rank* dapat menggunakan nilai *chi-square* yang akan dibandingkan dengan nilai *chi-square* tabel. Nilai *chi-square* pada pengujian ini sebesar 2,604 lebih kecil dari nilai *chi-square* tabel yang bernilai 3,841, artinya keputusan yang akan diambil gagal tolak  $H_0$ . Maka dapat disimpulkan artinya tidak ada perbedaan kurva *survival* antar kelompok pekerja dalam variabel anak.

### 4.3 Pemodelan Regresi Cox Proportional Hazard

Pemodelan analisis *survival* untuk data yang diberikan oleh PT. Taspen Kantor Cabang Utama Surabaya dapat dianalisis dengan menggunakan metode regresi *cox proportional hazard*. Langkah-langkah yang akan dilakukan dalam analisis ini adalah menguji asumsi *Proportional Hazard*, mengestimasi parameter, menguji signifikansi parameter, menghitung nilai *hazard ratio* dan menginterpretasi model regresi *cox proportional hazard*.

#### 4.3.1 Pengujian Asumsi Proportional Hazard

Dalam pemodelan regresi *cox proportional hazard* akan diuji asumsi *proportional hazard* dengan menggunakan uji *goodness of fit* pada *software* R. Pengujian asumsi *proportional hazard* akan diuji dengan nilai statistik uji dan *p-value*. Pengujian ini akan menggunakan taraf signifikansi 5%. Jika nilai statistik uji lebih besar daripada nilai *chi-square* tabel maka keputusan yang

diambil tolak  $H_0$ . Hal ini berlaku juga pada  $p$ -value lebih kecil dari taraf signifikan. Maka asumsi *proportional hazard* tidak terpenuhi.

**Tabel 4.10** Hasil *Goodness of Fit* Semua Variabel

Variabel	Statistik Uji	P-Value	Keputusan
Jenis Kelamin (1)	4,710	0,0300	Tolak $H_0$
Golongan Pekerja (3)	0,286	0,5931	Gagal Tolak $H_0$
Golongan Pekerja (4)	1,900	0,1680	Gagal Tolak $H_0$
Pasangan (1)	1,160	0,2822	Gagal Tolak $H_0$
Anak (1)	0,046	0,8306	Gagal Tolak $H_0$
Anak (2)	0,925	0,3361	Gagal Tolak $H_0$
Usia	$8.51.10^{-6}$	0,9977	Gagal Tolak $H_0$

Hasil keputusan dari pengujian *goodness of fit* pada Tabel 4.10 dengan menggunakan taraf signifikansi 5% adalah variabel jenis kelamin memiliki keputusan tolak  $H_0$  karena  $p$ -value lebih kecil dari taraf signifikan. Sedangkan  $p$ -value variabel lainnya lebih besar dari taraf signifikannya sehingga keputusan yang diambil adalah gagal tolak  $H_0$ . Begitu pula jika diuji dengan nilai *chi-square* tabel sebesar 3,841459, nilai statistik uji yang melebihi *chi-square* tabel hanya variabel jenis kelamin.

**Tabel 4.11** Hasil *Goodness of Fit* Tanpa Variabel Jenis Kelamin

Variabel	Statistik Uji	P-Value	Keputusan
Golongan (3)	0,418	0,5181	Gagal tolak $H_0$
Golongan (4)	2,939	0,0865	Gagal tolak $H_0$
Pasangan (1)	0,958	0,3277	Gagal tolak $H_0$
Anak (1)	0,600	0,4384	Gagal tolak $H_0$
Anak (2)	0,265	0,6065	Gagal tolak $H_0$
Usia	0,747	0,3874	Gagal tolak $H_0$

Tabel 4.11 merupakan hasil pengujian *proporsional hazard* tanpa menggunakan variabel jenis kelamin.  $P$ -value dari semua variabel dapat diambil keputusan gagal tolak  $H_0$  karena  $p$ -value melebihi taraf signifikan sehingga faktor golongan pekerja, pasangan, anak dan usia memenuhi asumsi *proporsional hazard*. Begitu pula jika diuji dengan menggunakan statistik uji, nilai

statistik uji lebih kecil daripada *chi-square* tabel sebesar 3,841459. Maka dapat disimpulkan bahwa faktor-faktor tersebut memenuhi asumsi *proportional hazard*.

#### 4.3.2 Estimasi Parameter Regresi Cox Proportional Hazard

Analisis selanjutnya adalah mengestimasi parameter yang signifikan. Variabel dependen yang akan digunakan adalah waktu *survival* pekerja dengan faktor-faktor yang diduga mempengaruhi variabel dependen tersebut yaitu jenis kelamin, pasangan, anak dan usia.

**Tabel 4.12** Estimasi Parameter Model *Cox Proportional Hazard*

Variabel	Parameter ( $\beta$ )	Exp ( $\beta$ )	Standar Error	Statistik Uji	P-Value
Golongan (3)	-0,05524	0,9463	0,3714	0,022121	0,8817
Golongan (4)	-0,7941	0,4520	0,4069	3,808684	0,0509
Pasangan (1)	-1,1880	0,3048	0,3323	12,78121	0,0003
Anak (1)	0,2475	1,2810	0,3505	0,498625	0,4801
Anak (2)	0,5865	1,7980	0,3348	3,068775	0,0798
Usia	$-9,375 \cdot 10^{-11}$	1,0000	$8,990 \cdot 10^{-11}$	1,087484	0,2969
<i>Likelihood Ratio</i>				23,57	0,00062

Tabel 4.12 dapat menunjukkan nilai estimasi parameternya sehingga didapatkan model *cox proportional hazard* sebagai berikut.

$$h(t) = h_0(t) \exp(-0,05524 X_{2(3)} - 0,7941 X_{2(4)} - 1,188 X_{3(1)} + 0,2475 X_{4(1)} + 0,5865 X_{4(2)} - 9,375 \cdot 10^{-11} X_5)$$

Berdasarkan model *cox proportional hazard* yang sudah terbentuk akan dilakukan uji serentak. Pengujian serentak dilakukan untuk mengetahui apakah variabel independen akan berpengaruh secara signifikan terhadap model. Uji ini akan menggunakan *p-value* dan statistik uji dari *likelihood ratio* pada Tabel 4.12. *P-value* dari *likelihood ratio* sebesar 0,03493. *P-value* ini akan dibandingkan dengan taraf signifikansi 5% apabila *p-value* kurang dari taraf signifikannya maka akan diperoleh keputusan tolak  $H_0$ . Dari



ketentuan tersebut keputusan yang diambil adalah tolak  $H_0$  karena  $p$ -value kurang dari 0,05 artinya minimal ada satu variabel independen yang berpengaruh secara signifikan terhadap model. Selain dengan  $p$ -value pengambilan keputusan dalam pengujian serentak bisa dengan cara menghitung nilai statistik uji atau  $G^2$  yang akan dibandingkan dengan nilai  $chi$ -square tabel. Pada Tabel 4.12 nilai statistik ujinya sebesar 23,57 dengan nilai  $chi$ -square tabel sebesar 1,635383. Dari nilai tersebut didapatkan hasil bahwa nilai statistik uji lebih besar daripada nilai  $chi$ -square tabel maka keputusan yang diambil adalah tolak  $H_0$ . Maka dapat disimpulkan model yang terbentuk secara bersama-sama telah berpengaruh secara signifikan.

Selanjutnya akan dilakukan uji parsial. Pengujian ini akan menggunakan  $p$ -value dan nilai statistik uji dari setiap parameter.  $P$ -value ini akan dibandingkan dengan taraf signifikansi 5% apabila  $p$ -value lebih kecil dari  $\alpha$  maka akan diambil keputusan tolak  $H_0$ . Selain itu akan dilakukan perbandingan nilai statistik uji dengan  $chi$ -square tabel, apabila nilai dari  $chi$ -square tabel lebih besar dari statistik uji maka akan diambil keputusan tolak  $H_0$ . Dari Tabel 4.4 dapat dilihat bahwa  $p$ -value dari variabel pasangan sebesar 0,000349 lebih kecil dari  $\alpha$ . Hal ini mengakibatkan keputusan yang diambil tolak  $H_0$ . Hal ini berlaku pada perbandingan nilai statistik uji sebesar 12,78122 dengan nilai  $chi$ -square tabel sebesar 3,841459 karena nilai  $chi$ -square tabel lebih kecil dari nilai statistik uji. Maka diperlukan eliminasi *backward* untuk menentukan model *cox proportional hazard* karena banyak parameter yang tidak signifikan.

#### 4.3.3 Seleksi Model Terbaik Regresi Cox Proportional Hazard

Seleksi model terbaik dilakukan dengan menggunakan eliminasi *backward* dimana variabel yang tidak signifikan dibuang satu persatu. Hal ini dilakukan untuk mendapatkan variabel yang signifikan. Proses ini dilakukan dengan menggunakan nilai *Akaike's information Criterion* (AIC) terkecil. Berdasarkan hasil dari Lampiran 5 nilai AIC terkecil sebesar 441,16. Model yang terbentuk dari variabel golongan dan pasangan sehingga dilakukan

estimasi parameter kembali dari variabel yang signifikan. Berikut ini hasil estimasi parameter dengan variabel yang signifikan.

**Tabel 4.13** Estimasi Parameter Model *Cox Proportional Hazard* Terbaik

Variabel	Parameter ( $\beta$ )	Exp ( $\beta$ )	Standar Error	Statistik Uji	P-Value
Golongan (3)	-0,1934	0,8241	0,3628	0,28417	0,59398
Golongan (4)	-1,0292	0,3573	0,3899	6,96775	0,00830
Pasangan (1)	-1,0651	0,3447	0,3065	12,07591	0,00051
<i>Likelihood Ratio</i>				19,78	0,00019

Tabel 4.13 dapat menunjukkan nilai estimasi parameternya setelah diseleksi sehingga didapatkan model *cox proportional hazard* terbaik sebagai berikut.

$$h(t) = h_0(t) \exp(-0,1934X_{2(3)} - 1,0292X_{2(4)} - 1,0651X_{3(1)})$$

Berdasarkan model *cox proportional hazard* yang sudah terbentuk akan dilakukan uji serentak. Pengujian serentak dilakukan untuk mengetahui apakah variabel independen akan berpengaruh secara signifikan terhadap model terbaik. Uji ini akan menggunakan *p-value* dan statistik uji dari *likelihood ratio* pada Tabel 4.13. *P-value* dari *likelihood ratio* sebesar 0,00019. *P-value* ini akan dibandingkan dengan taraf signifikansi 5% apabila *p-value* kurang dari taraf signifikannya maka akan diperoleh keputusan tolak  $H_0$ . Dari ketentuan tersebut keputusan yang diambil adalah tolak  $H_0$  karena *p-value* kurang dari 0,05 artinya minimal ada satu variabel independen yang berpengaruh secara signifikan terhadap model. Selain dengan *p-value* pengambilan keputusan dalam pengujian serentak bisa dengan cara menghitung nilai statistik uji atau  $G^2$  yang akan dibandingkan dengan nilai *chi-square* tabel. Pada Tabel 4.13 nilai statistik ujinya sebesar 19,78 dengan nilai *chi-square* tabel sebesar 0,351846. Dari nilai tersebut didapatkan hasil bahwa nilai statistik uji lebih besar daripada nilai *chi-square* tabel maka keputusan yang diambil adalah tolak  $H_0$ . Maka dapat disimpulkan

model yang terbentuk secara bersama-sama telah berpengaruh secara signifikan.

Selanjutnya akan dilakukan uji parsial. Pengujian ini akan menggunakan *p-value* dan nilai statistik uji dari setiap parameter. *P-value* ini akan dibandingkan dengan taraf signifikansi 5% apabila *p-value* lebih kecil dari  $\alpha$  maka akan diambil keputusan tolak  $H_0$ . Selain itu akan dilakukan perbandingan nilai statistik uji dengan *chi-square* tabel, apabila nilai dari *chi-square* tabel lebih besar dari statistik uji maka akan diambil keputusan tolak  $H_0$ . Dari Tabel 4.4 dapat dilihat bahwa *p-value* dari variabel golongan 4 dan pasangan berturut-turut sebesar 0,0083 dan 0,00051 lebih kecil dari  $\alpha$ . Hal ini mengakibatkan keputusan yang diambil tolak  $H_0$ . Hal ini berlaku pada perbandingan nilai statistik uji berturut-turut sebesar 6,96775 dan 12,07591 dengan nilai *chi-square* tabel sebesar 3,841459 karena nilai *chi-square* tabel lebih kecil maka dapat disimpulkan variabel golongan 4 dan pasangan berpengaruh secara signifikan terhadap waktu *survival* pekerja. Sedangkan variabel lainnya tidak berpengaruh secara signifikan. Hal ini dapat dikarenakan *human error* seperti kesalahan *input* data.

#### 4.3.4 Interpretasi Model Cox Proportional Hazard

Model *cox proportional hazard* diperoleh dengan melalui beberapa tahap. Hal ini diperlukan untuk mengetahui variabel-variabel mana saja yang berpengaruh secara signifikan terhadap waktu *survival* pekerja. Variabel golongan 4 dan pasangan berpengaruh secara signifikan terhadap model sehingga dapat diinterpretasikan dengan nilai *odds ratio*. Nilai *odds ratio* didapatkan dengan menggunakan nilai *exponential* dari koefisien pada model. Nilai *exponential* dari koefisien golongan 4 dan pasangan dapat dilihat pada Tabel 4.13 dengan nilai berturut-turut sebesar 0,3573 dan 0,3447. Nilai tersebut dapat diinterpretasikan menjadi pekerja dengan golongan 4 memiliki resiko 0,3573 kali untuk meninggal daripada pekerja dengan golongan 1 dan 2. Selain interpretasi dari variabel golongan, selanjutnya akan dilakukan interpretasi variabel pasangan. Pekerja yang mempunyai 1 pasangan memiliki resiko 0,3447 kali untuk meninggal daripada

pekerja yang tidak memiliki pasangan atau dengan kata lain pekerja yang memiliki 1 pasangan memiliki kecenderungan lebih kecil untuk meninggal daripada pekerja yang tidak memiliki pasangan. Berdasarkan jurnal dari *Emory University School of Medicine* para janda atau duda memiliki resiko lebih tinggi meninggal daripada yang yang berpasangan karena orang-orang yang memiliki pasangan cenderung saling menjaga pasangannya dan saling mengingatkan kesehatan (Quyyumi, 2017).

#### 4.3.5 Analisis Regresi Cox Proportional Hazard dengan Menggunakan Metode Bootstrap

Proses *bootstrap* dalam analisis ini dilakukan dengan cara mengambil acak sampel data baru dengan menggunakan data asli secara berpasangan. Dalam penelitian ini, sampel baru akan diulang sebanyak 5000 kali. Kemudian hasil perulangan tersebut akan dicari parameter untuk setiap perulangan dengan menggunakan regresi *cox proportional hazard*. Parameter-parameter yang sudah didapatkan akan dirata-rata dan dicari *confidence interval*-nya.

**Tabel 4.14** *Confidence Interval 95% dari Parameter dengan Metode Bootstrap*

Variabel	Confidence Interval	Parameter ( $\beta$ )	Exp ( $\beta$ )
Golongan (3)	[-0,20;-0,18]	-0,19	0,826
Golongan (4)	[-1,13;-1,10]	-1,11	0,329
Pasangan (1)	[-1,19;-1,17]	-1,18	0,307

Uji signifikansi parameter pada metode *bootstrap* menggunakan nilai *confidence interval* apabila nilai *confidence interval* memuat nilai nol maka parameter tersebut tidak signifikan. Namun apabila nilai *confidence interval* tidak memuat nilai nol maka parameter tersebut signifikan. Tabel 4.14 menunjukkan bahwa semua parameter tidak memuat nilai nol sehingga dapat disimpulkan bahwa semua parameter signifikan. Model yang terbentuk dari parameter-parameter tersebut adalah sebagai berikut.

$$h(t) = h_0(t) \exp(-0,19X_{2(3)} - 1,11X_{2(4)} - 1,18X_{3(1)})$$

Berdasarkan model tersebut faktor-faktor yang diduga mempengaruhi kematian pekerja dapat diinterpretasikan dengan menggunakan nilai eksponen dari koefisien setiap variabel pada Tabel 4.14. Pekerja dengan golongan 3 memiliki resiko 0,826 kali untuk meninggal daripada pekerja dengan golongan 1&2. Selain itu pekerja dengan golongan 4 memiliki resiko 0,329 kali untuk meninggal daripada golongan 1&2. Variabel pasangan memiliki nilai eksponen koefisien sebesar 0,307 sehingga dapat diinterpretasikan bahwa pekerja yang tidak mempunyai pasangan memiliki resiko 0,307 kali untuk meninggal daripada pekerja yang mempunyai satu pasangan. Dalam metode ini hasil parameter yang signifikan cocok dengan data yang diberikan oleh PT.Taspen Kantor Cabang Utama Surabaya.

*“Halaman ini sengaja dikosongkan”*

## **BAB V**

### **KESIMPULAN DAN SARAN**

#### **5.1 Kesimpulan**

Setelah melalui semua tahapan analisis mengenai data *survival* pekerja maka dapat disimpulkan sebagai berikut.

1. Hasil karakteristik data pekerja adalah pekerja mampu bertahan hidup hingga berumur 63 tahun dengan rata-rata waktu *survival* pekerja adalah 24 tahun. Jumlah data tersensor pada data pekerja sebanyak 40 orang dan data tidak tersensor sebanyak 60 orang. Jenis kelamin pria pada data pekerja memiliki jumlah paling banyak daripada wanita. Golongan pekerja terbanyak dari 100 orang pekerja adalah golongan 3. Pekerja yang memiliki 1 pasangan lebih banyak jumlahnya daripada pekerja yang tidak memiliki pasangan dari 100 orang data pekerja. Pekerja yang tidak memiliki anak mempunyai jumlah terbesar dibanding pekerja yang memiliki 1 anak maupun 2 anak dari data pekerja.
2. Hasil analisis karakteristik kurva Kaplan-Meier untuk semua variabel menurun sampai pada rentang waktu 40 tahun. Tingkat kematian paling tinggi pada variabel jenis kelamin, golongan, pasangan dan anak berturut-turut adalah pekerja jenis kelamin wanita, pekerja pada golongan 2, pekerja yang tidak memiliki pasangan dan pekerja yang memiliki 2 anak. Berdasarkan uji *log-rank* pada variabel golongan pekerja dan pasangan memiliki perbedaan waktu ketahanan hidup antar kelompok dalam variabelnya. Sedangkan kurva *survival* antar kelompok dalam variabel jenis kelamin dan anak tidak memiliki perbedaan.
3. Faktor yang berpengaruh terhadap model regresi *cox proportional hazard* adalah variabel golongan dan pasangan berdasarkan hasil seleksi variabel. Hal ini menunjukkan bahwa variabel yang berpengaruh terhadap model yang dibangun adalah variabel golongan dan pasangan. Selanjutnya dilakukan *resampling bootstrap* untuk memperlihatkan data keterwakilannya. Dalam pengujian

regresi *cox proportional hazard* dengan metode *bootstrap* didapati semua parameter signifikan. Hal ini menunjukkan bahwa dari model yang dibangun dengan menggunakan simulasi *bootstrap*, variabel yang berpengaruh terhadap model ini adalah golongan pekerja dan pasangan.

## **5.2 Saran**

Bagi peneliti selanjutnya, diharapkan pada penelitian berikutnya dapat menambahkan faktor-faktor lainnya yang mempengaruhi ketahanan hidup pekerja. Hal ini diharapkan dapat memberi pembaca pengetahuan lebih luas tentang faktor-faktor yang mempengaruhi ketahanan hidup pekerja. Selain itu dalam mengestimasi parameter dengan *bootstrap* tampilkan bukti yang memastikan proporsi sesuai dengan data asli.



## DAFTAR PUSTAKA

- Anindita, A. (2017). *Perhitungan Hak dan Manfaat yang Diterima Ahli Waris Peserta Meninggal Aktif pada PT. Taspen (Persero) KC Bandar Lampung*. Bandar Lampung: Politeknik Negeri Lampung.
- Bennet, P. J. (2009). Introduction to the Bootstrap and Robust Statistics. *Winter Term*, 711/712.
- Chiu, Rahman, & Vigod. (2018). Mortality in single fathers compared with single mothers and partnered parents: a population-based cohort study. *Lancet Public Health*, 2468-2667.
- Collet, D. (2003). *Modeling Survival Data in Medical Research*. London: Chapman & Hall/CRC.
- Cox, D. R. (1972). Regression Models and Life Tables (with Discussion). *Journal of the Royal Statistical Society*.
- Dister, N. S. (1982). *Pengantar Psikologi Agama : Pengalaman dan Motivasi Beragama*. Jakarta: Leppenas.
- Gonzalez-Mule, E., & Cockburn, B. (2016). Worked to Death: The Relationships of Job Demands and Job Control with Mortality. *Personnel Psychology*, Volume 70, Issue 1.
- Hartini, N. (2007). Deat and Dying. *Sress Management dalam Berbagai Kehidupan*. Bandung.
- Hasan, A. W. (2006). *Aplikasi Strategi Dan Model Kecerdasan Spiritual (SQ) Rasulullulah diMasa Kini*. Yogyakarta: IRCiSod.
- Hosmer, D. W., Lemeshow, S., & May, S. (2008). *Applied Survival Analysis : Regression Modelling of Time Event Data*. New Jersey: John Wiley.
- Ihwah, A. (2012). *Bootstrap pada Data Tersensor untuk Estimasi Beta dalam Model Regresi Cox*. Yogyakarta: Gadjah Mada.
- Klein, K. a. (2012). *Survival Analysis*. New York: Springer.
- Kleinbaum, & Klein. (2005). *Survival Analysis*. London: Springer Science Business Media, Inc.
- Le, C. T. (1997). *Applied Survival Analysis*. New York: John Wiley and Sons, Inc.

- Lee, E. T. (1980). *Statistical Methods for Survival Data Analysis*. California: Departement of Biostatistics and Epidemiology School of Public Health University of Oklahoma, Oklahoma City, Oklahoma. Life Learning Publications Belmont.
- Lee, E., & Wang, J. (2003). *Statistical Methods for Survival Data Analysis. Ed ke-3*. New York: John Wiley & Son, Inc.
- Lee, E., & Wang, J. (2003). *Statistical Methods for Survival Data Analysis Third Edition*. New Jersey: John Wiley & Sons.
- Quyyumi, A. (2017). Marital Status and Outcomes in Patients With Cardiovascular Disease. *Journal of the American Heart Association*.
- Rahayu, N., Setiawan, A., & Mahatma, T. (2012, Oktober 6). Analisis Regresi Cox Proportional Hazard pada Ketahanan Hidup Pasien Diabetes Mellitus. *Prosiding Seminar Nasional Matematika UNS*.
- Sahinler, S., & Topuz, D. (2007). Bootstrap and Jackknife Resampling Algorithms for Estimation of Regression Parameters. *JAQM no. 2*, 188-199.
- Santrock, J. W. (2002). *Perkembangan Masa Hidup (edisi kelima)*. Jakarta: Erlangga.

**Lampiran 1 Data Pekerja Pensiun Meninggal Aktif**

T	D	X1	X2	X3	X4	X5
20,93425	0	1	3	1	2	52,50685
35,90137	0	1	3	1	0	62,53151
8,846575	1	1	3	1	2	46,28493
11,01096	0	0	2	0	0	43,39726
34,89589	0	1	0	1	1	56,56164
22,48493	1	0	0	0	0	52,91781
31,94521	0	1	0	0	0	56,93425
11,06027	0	0	2	1	2	51,40274
11,03288	0	1	2	1	2	53,98356
34,9726	0	1	0	1	2	59,89863
31,32603	1	1	3	1	1	55,38356
27,0274	1	1	2	1	2	51,00822
32,16164	1	0	3	1	1	53,29589
11,04932	0	1	2	1	1	48,66849
16,04384	1	1	3	1	1	45,1863
⋮	⋮	⋮	⋮	⋮	⋮	⋮
22,52877	1	0	0	0	0	51,48767
29,78082	1	1	2	1	2	55,11233
30,88493	0	1	0	1	2	52,42466
12,33425	1	0	3	0	0	48,0137
32,91233	0	1	0	1	1	54,49589
30,45753	1	1	2	1	2	53,8137

## Lampiran 2 Output SPSS Uji Log-Rank Jenis Kelamin

a. Jenis kelamin

### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	,374	1	,541

Test of equality of survival distributions for the different levels of X1.

b. Golongan Pekerja

### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	33,428	3	,000

Test of equality of survival distributions for the different levels of X2.

c. Pasangan

### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	8,077	1	,004

Test of equality of survival distributions for the different levels of X3.

d. Anak

### Overall Comparisons

	Chi-Square	df	Sig.
Log Rank (Mantel-Cox)	2,604	2	,061

Test of equality of survival distributions for the different levels of X4.

### Lampiran 3 *Output R Goodness of Fit*

a. Semua variabel

```
> test.ph <- cox.zph(modelcox, transform="rank", global=TRUE)
> test.ph
```

	rho	chisq	p
X11	-0.229228	4.71e+00	0.0300
X23	-0.072972	2.86e-01	0.5931
X24	0.192756	1.90e+00	0.1680
X31	0.137671	1.16e+00	0.2822
X41	-0.024904	4.58e-02	0.8306
X42	0.131270	9.25e-01	0.3361
X5	-0.000352	8.51e-06	0.9977
GLOBAL	NA	1.50e+01	0.0354

b. Tanpa variabel jenis kelamin

```
> test.ph <- cox.zph(modelcox, transform="rank", global=TRUE)
> test.ph
```

	rho	chisq	p
X22	-0.240	2.939	0.08648
X23	-0.408	9.331	0.00225
X31	0.123	0.958	0.32765
X41	0.103	0.600	0.43845
X42	0.071	0.265	0.60655
X5	-0.115	0.747	0.38740
GLOBAL	NA	10.638	0.10024

**Lampiran 4 Estimasi Parameter Seluruh Variabel yang Signifikan**

```

> model.A <- coxph(Surv(Data_Baru$T, Data_Baru$D==1)~X2+X3
+X4+X5, data=Data_Baru,method = "breslow")
> summary(model.A)
coxph(formula = Surv(Data_Baru$T, Data_Baru$D == 1) ~ X2 +
  X3 +
    X4 + X5, data = Data_Baru, method = "breslow")

n= 100, number of events= 60

              coef exp(coef)    se(coef)      z Pr(>|z|)
X23 -5.524e-02  9.463e-01  3.714e-01 -0.149 0.881766
X24 -7.941e-01  4.520e-01  4.069e-01 -1.952 0.050969 .
X31 -1.188e+00  3.048e-01  3.323e-01 -3.576 0.000349 ***
X41  2.475e-01  1.281e+00  3.505e-01  0.706 0.480159
X42  5.865e-01  1.798e+00  3.348e-01  1.752 0.079806 .
X5   -9.375e-11 1.000e+00  8.990e-11 -1.043 0.296990
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 '
' 1

              exp(coef) exp(-coef) lower .95 upper .95
X23      0.9463      1.0568      0.4569      1.9596
X24      0.4520      2.2125      0.2036      1.0034
X31      0.3048      3.2807      0.1589      0.5846
X41      1.2808      0.7808      0.6444      2.5457
X42      1.7977      0.5563      0.9327      3.4651
X5        1.0000      1.0000      1.0000      1.0000

Concordance= 0.72 (se = 0.042 )
Rsquare= 0.21 (max possible= 0.989 )
Likelihood ratio test= 23.57 on 6 df, p=0.0006251
Wald test               = 19.23 on 6 df, p=0.003786
Score (logrank) test = 25.43 on 6 df, p=0.0002844

```

### Lampiran 5. Seleksi Model Terbaik dengan Kriteria AIC

```
MT=step(model.A,direction="backward",criterion="AIC")
Start:  AIC=443.36
Surv(Data_Baru$T, Data_Baru$D == 1) ~ x2 + x3 + x4 + x5
```

	Df	AIC
- x5	1	442.37
- x4	2	442.40
<none>		443.36
- x2	2	446.30
- x3	1	452.73

```
Step:  AIC=442.37
Surv(Data_Baru$T, Data_Baru$D == 1) ~ x2 + x3 + x4
```

	Df	AIC
- x4	2	441.16
<none>		442.37
- x2	2	447.30
- x3	1	452.82

```
Step:  AIC=441.16
Surv(Data_Baru$T, Data_Baru$D == 1) ~ x2 + x3
```

	Df	AIC
<none>		441.16
- x2	2	448.03
- x3	1	449.52

## Lampiran 6. Estimasi Parameter Model Terbaik untuk Regresi Cox Proportional Hazard

```
> model.A <- coxph(Surv(Data_Baru$T, Data_Baru$D==1)~X2+X3
, data=Data_Baru,method = "breslow")
> summary(model.A)
```

Call:  
coxph(formula = Surv(Data\_Baru\$T, Data\_Baru\$D == 1) ~ X2 +  
X3,  
data = Data\_Baru, method = "breslow")

n= 100, number of events= 60

	coef	exp(coef)	se(coef)	z	Pr(> z )
X23	-0.1934	0.8241	0.3628	-0.533	0.59398
X24	-1.0292	0.3573	0.3899	-2.639	0.00830 **
X31	-1.0651	0.3447	0.3065	-3.475	0.00051 ***

---  
Signif. codes:  
0 '\*\*\*' 0.001 '\*\*' 0.01 '\*' 0.05 '.' 0.1 ' ' 1

	exp(coef)	exp(-coef)	lower .95	upper .95
X23	0.8241	1.213	0.4047	1.6782
X24	0.3573	2.799	0.1664	0.7672
X31	0.3447	2.901	0.1890	0.6285

Concordance= 0.682 (se = 0.04 )  
Rsquare= 0.179 (max possible= 0.989 )  
Likelihood ratio test= 19.78 on 3 df, p=0.0001886  
Wald test = 20.4 on 3 df, p=0.0001405  
Score (logrank) test = 21.77 on 3 df, p=7.272e-05



**Lampiran 7 Syntax Regresi Cox *Proportional Hazard* dengan Metode *Bootstrap***

```
coxboot <- function(Data_Baru,B)
{
  has <- matrix(0,B,3)
  for (i in 1:B)
  {
    sampelbaru<-sample(1:100,100,replace=TRUE)
    data <- Data_Baru[sampelbaru,]
    Data_Baru.reg<- coxph(Surv(T, D) ~ x2+x3+x4+x5, data)
    has[i,] <- Data_Baru.reg[[1]]
  }
  has1 <- has
  for (i in 1:3)
  {
    has1[,i] <- sort(has[,i])
  }
  return(has1)
}
ulang <-coxboot(Data_Baru,5000)
ci.mean(ulang[,1])
ci.mean(ulang[,2])
ci.mean(ulang[,3])
ci.mean(ulang[,4])
ci.mean(ulang[,5])
ci.mean(ulang[,6])
```

## Lampiran 8 Surat Keterangan Legalitas Data



PT. TASPEN (PERSERO)  
KANTOR CABANG UTAMA SURABAYA  
Jl. Raya Diponegoro 193 Surabaya 60421  
Telp. Hunting (031) 5676356 Fax. (031) 5671029

### SURAT KETERANGAN

Nomor : KET-09/CU.05.3/062018

Yang bertanda tangan di bawah ini menerangkan bahwa :

1. Mahasiswa Statistika FMKSD-ITS dengan identitas berikut :

N a m a	: Nadia Rahma Nur Insyira
N R P	: 0621144000080 (1314 100 080)
Program Studi	: Sarjana (S1)
Departemen	: Statistika FMKSD

Telah mengambil data di Instansi kami pada Bidang Layanan dan Manfaat Seksi Kepesertaan untuk keperluan Tugas Akhir.

2. Tidak keberatan nama Instansi dicantumkan dalam Tugas Akhir mahasiswa Statistika yang akan disimpan di Perpustakaan ITS dan dibaca di lingkungan ITS.
3. Tidak keberatan bahwa hasil analisis data dari Instansi dipublikasikan dalam E-Journal ITS yaitu Journal Sains dan Seni ITS.

Surabaya, 28 Juni 2018  
PT. TASPEN (PERSERO)  
Wakil Kepala Kantor Cabang Utama

  
TAMSIR  
NIK. 198819651546

## BIODATA PENULIS



Penulis bernama Nadia Rahma Nur Insyira dilahirkan di Kota Jakarta pada tanggal 30 September 1995 sebagai anak pertama dari dua bersaudara. Penulis menempuh pendidikan formal di SD Al-Hikmah Surabaya, SMP Negeri 12 Surabaya, dan SMA Negeri 15 Surabaya. Kemudian penulis diterima sebagai Mahasiswa departemen Statistika FMKSD ITS melalui jalur SBMPTN pada tahun 2014 dengan NRP 06211440000080. Selama masa perkuliahan, penulis memiliki pengalaman dalam berpartisipasi di berbagai kepanitiaan diantaranya sebagai GERIGI ITS pada tahun 2015 dan 2016, Anggota Divisi *Sponsorship* dalam *event* tahunan PRS (Pekan Raya Statistika) 2016, Anggota BEM Divisi BSO VIVAT Press, Sekretaris Pers HIMASTA-ITS. Penulis juga pernah mengikuti pelatihan yaitu LKMM-Pra-TD. Apabila pembaca ingin memberi kritik dan saran serta diskusi lebih lanjut mengenai Tugas Akhir ini, dapat menghubungi penulis melalui email [narasyira@gmail.com](mailto:narasyira@gmail.com).

*“Halaman ini sengaja dikosongkan”*